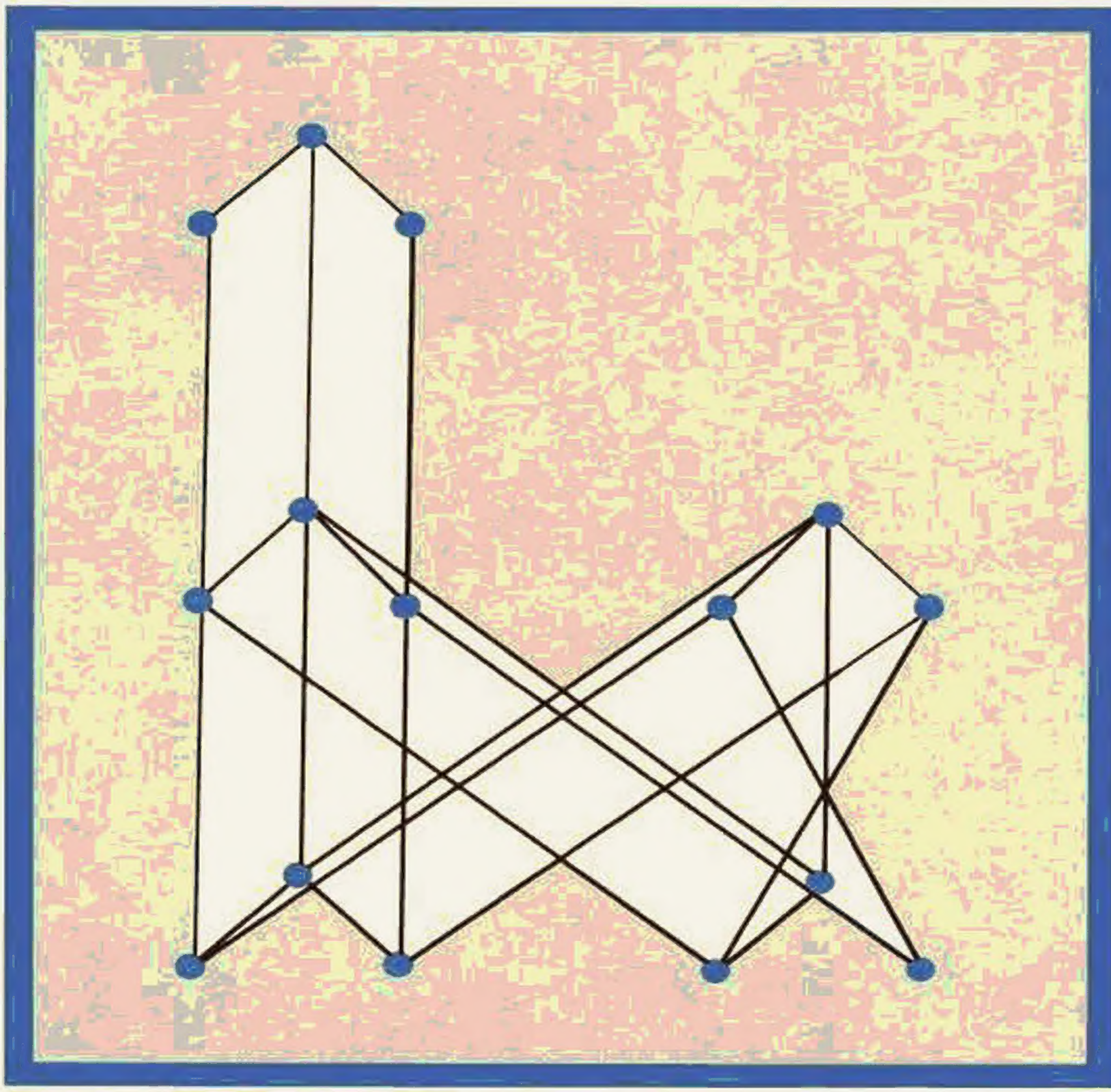
ENUMERATIVE COMBINATORICS
VOLUME II

RICHARD P. STANLEY



Enumerative Combinatorics

This is the second of a two-volume basic introduction to enumerative combinatorics at a level suitable for graduate students and research mathematicians.

This volume covers the composition of generating functions, trees, algebraic generating functions, D-finite generating functions, noncommutative generating functions, and symmetric functions. The chapter on symmetric functions provides the only available treatment of this subject suitable for an introductory graduate course and focusing on combinatorics, especially the Robinson–Schensted–Knuth algorithm. Also covered are connections between symmetric functions and representation theory. An appendix (written by Sergey Fomin) covers some deeper aspects of symmetric function theory, including jeu de taquin and the Littlewood–Richardson rule.

As in Volume 1, the exercises play a vital role in developing the material. There are over 250 exercises, all with solutions or references to solutions, many of which concern previously unpublished results.

Graduate students and research mathematicians who wish to apply combinatorics to their work will find this an authoritative reference.

Richard P. Stanley is Professor of Applied Mathematics at the Massachusetts Institute of Technology. He has held visiting positions at UCSD, the University of Strasbourg, California Institute of Technology, the University of Augsburg, Tokai University, and the Royal Institute of Technology in Stockholm. He has published over 100 research papers in algebraic combinatorics. In addition to the two-volume *Enumerative Combinatorics*, he has published one other book, *Combinatorics and Commutative Algebra* (Birkhäuser; second edition, 1997). He is a fellow of the American Academy of Arts and Sciences, a member of the National Academy of Sciences, and a recipient of the Pólya Prize in Applied Combinatorics awarded by the Society for Industrial and Applied Mathematics.

CAMBRIDGE STUDIES IN ADVANCED MATHEMATICS 62

Already published

2 K. Petersen *Ergodic theory*
3 P.T. Johnstone *Stone spaces*
4 W.H. Schikhof *Ultrametric calculus*
5 J.-P. Kahane *Some random series of functions, 2nd edition*
7 J. Lambek & P.J. Scott *Introduction to higher-order categorical logic*
8 H. Matsumura *Commutative ring theory*
9 C.B. Thomas *Characteristic classes and the cohomology of finite groups*
10 M. Aschbacher *Finite group theory*
11 J.L. Alperin *Local representation theory*
12 P. Koosis *The logarithmic integral I*
13 A. Pietsch *Eigenvalues and s-numbers*
14 S.J. Patterson *An introduction to the theory of the Riemann zeta-function*
15 H.J. Baues *Algebraic homotopy*
16 V.S. Varadarajan *Introduction to harmonic analysis on semisimple Lie groups*
17 W. Dicks & M. Dunwoody *Groups acting on graphs*
18 L.J. Corwin & F.P. Greenleaf *Representations of nilpotent Lie groups and their applications*
19 R. Fritsch & R. Piccinini *Cellular structures in topology*
20 H. Klingen *Introductory lectures on Siegel modular forms*
21 P. Koosis *The logarithmic integral II*
22 M.J. Collins *Representations and characters of finite groups*
24 H. Kunita *Stochastic flows and stochastic differential equations*
25 P. Wojtaszczyk *Banach spaces for analysts*
26 J.E. Gilbert & M.A.M. Murray *Clifford algebras and Dirac operators in harmonic analysis*
27 A. Frohlich & M.J. Taylor *Algebraic number theory*
28 K. Goebel & W.A. Kirk *Topics in metric fixed point theory*
29 J.F. Humphreys *Reflection groups and Coxeter groups*
30 D.J. Benson *Representations and cohomology I*
31 D.J. Benson *Representations and cohomology II*
32 C. Allday & V. Puppe *Cohomological methods in transformation groups*
33 C. Soule *et al.* *Lectures on Arakelov geometry*
34 A. Ambrosetti & G. Prodi *A primer of nonlinear analysis*
35 J. Palis & F. Takens *Hyperbolicity, stability and chaos at homoclinic bifurcations*
37 Y. Meyer *Wavelets and operators 1*
38 C. Weibel *An Introduction to homological algebra*
39 W. Bruns & J. Herzog *Cohen–Macaulay rings*
40 V. Snaith *Explicit Brauer induction*
41 G. Laumon *Cohomology of Drinfeld modular varieties: Part I*
42 E.B. Davies *Spectral theory and differential operators*
43 J. Diestel, H. Jarchow & A. Tonge *Absolutely summing operators*
44 P. Mattila *Geometry of sets and measures in Euclidean spaces*
45 R. Pinsky *Positive harmonic functions and diffusion*
46 G. Ienenbaum *Introduction to analytic and probabilistic number theory*
47 C. Peskine *Complex projective geometry*
48 Y. Meyer & R. Coifman *Wavelets*
49 R. Stanley *Enumerative combinatorics I*
50 I. Porteous *Clifford algebras and the classical groups*
51 M. Audin *Spinning tops*
52 V. Jurdjevic *Geometric control theory*
53 H. Volklein *Groups as Galois groups*
54 J. Le Potier *Lectures on vector bundles*
55 D. Bump *Automorphic forms and representations*
56 G. Laumon *Cohomology of Drinfeld modular varieties II*
60 M.P. Brodmann & R.Y. Sharp *Local cohomology*

ENUMERATIVE COMBINATORICS

Volume 2

RICHARD P. STANLEY
Massachusetts Institute of Technology

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 1999
First paperback edition 2001

Printed in the United States of America

Typeset in Times Roman 10/12 pt. in LaTeX 2_ε [TB]

A catalog record for this book is available from the British Library

Library of Congress Cataloging in Publication data is available

ISBN 0 521 56069 1 hardback
ISBN 0 521 78987 7 paperback

Foreword

Most textbooks written in our day have a short half-life. Published to meet the demands of a lucrative but volatile market, inspired by the table of contents of some out-of-print classic, garnished with multicolored tables, enhanced by nutshell summaries, enriched by exercises of dubious applicability, they decorate the shelves of college bookstores come September. The leftovers after Registration Day will be shredded by Christmas, unwanted even by remainder bookstores. The pageant is repeated every year, with new textbooks on the same shelves by other authors (or a new edition if the author is the same), as similar to the preceding as one can make them, short of running into copyright problems.

Every once in a long while, a textbook worthy of the name comes along; invariably, it is likely to prove *aere perennius*: Weber, Bertini, van der Waerden, Feller, Dunford and Schwartz, Ahlfors, Stanley.

The mathematical community professes a snobbish distaste for expository writing, but the facts are at variance with the words. In actual reality, the names of authors of the handful of successful textbooks written in this century are included in the list of the most celebrated mathematicians of our time.

Only another textbook writer knows the pains and the endless effort that goes into this kind of writing. The amount of time that goes into drafting a satisfactory exposition is always underestimated by the reader. The time required to complete one single chapter exceeds the time required to publish a research paper. But far from wasting his or her time, the author of a successful textbook will be amply rewarded by a renown that will spill into the distant future. History is more likely to remember the name of the author of a definitive exposition than the names of many a research mathematician.

I find it impossible to predict when Richard Stanley's two-volume exposition of combinatorics may be superseded. No one will dare try, let alone be able, to match the thoroughness of coverage, the care for detail, the definitiveness of proof, the elegance of presentation. Stanley's book possesses that rarest quality among textbooks: you can open it at any page and start reading with interest without having to hark back to page one for previous explanations.

Combinatorics, which only thirty years ago was a fledgling among giants, may well be turning out to be a greater giant, thanks largely to Richard Stanley's work. Every one who deals with discrete mathematics, from category theorists to molecular biologists, owes him a large debt of gratitude.

Gian-Carlo Rota
March 21, 1998

Preface

This is the second (and final) volume of a graduate-level introduction to enumerative combinatorics. To those who have been waiting twelve years since the publication of Volume 1, I can only say that no one is more pleased to see Volume 2 finally completed than myself. I have tried to cover what I feel are the fundamental topics in enumerative combinatorics, and the ones that are the most useful in applications outside of combinatorics. Though the book is primarily intended to be a textbook for graduate students and a resource for professional mathematicians, I hope that undergraduates and even bright high-school students will find something of interest. For instance, many of the 66 combinatorial interpretations of Catalan numbers provided by Exercise 6.19 should be accessible to undergraduates with a little knowledge of combinatorics.

Much of the material in this book has never appeared before in textbook form. This is especially true of the treatment of symmetric functions in Chapter 7. Although the theory of symmetric functions and its connections with combinatorics is in my opinion one of the most beautiful topics in all of mathematics, it is a difficult subject for beginners to learn. The superb book by Macdonald on symmetric functions is highly algebraic and eschews the fundamental combinatorial tool in this subject, viz., the Robinson–Schensted–Knuth algorithm. I hope that Chapter 7 adequately fills this gap in the mathematical literature. Chapter 7 should be regarded as only an introduction to the theory of symmetric functions, and not as a comprehensive treatment.

As in Volume 1, the exercises play a vital role in developing the subject. If in reading the text the reader is left with the feeling of "what's it good for?" and is not satisfied with the applications presented there, then (s)he should turn to the exercises. Thanks to the wonders of electronic word processing, I found it much easier than for Volume 1 to assemble a wide variety of exercises and solutions.

I am grateful to the many persons who have contributed in a number of ways to the improvement of this book. Special thanks go to Sergey Fomin for his many suggestions related to Chapter 7, and especially for his masterful exposition of the difficult material of Appendix 1. Other persons who have carefully read portions of earlier versions of the book and who have offered valuable suggestions

are Christine Bessenrodt, Francesco Brenti, Persi Diaconis, Wungkum Fong, Phil Hanlon, and Michelle Wachs. Robert Becker typed most of Chapter 5, and Tom Roby and Bonnie Friedman provided invaluable TEX assistance. The following persons at Cambridge University Press and TechBooks have been a pleasure to work with throughout the writing and production of this book: Catherine Felgar, Shamus McGillicuddy, Andrew Wilson, and especially Lauren Cowles, whose patience and support is greatly appreciated. The following additional persons have made at least one significant contribution that is not explicitly mentioned in the text, and I regret if I have inadvertently omitted anyone else who belongs on this list: Christos Athanasiadis, Anders Björner, Mireille Bousquet-Mélou, Bradley Brock, David Buchsbaum, Emeric Deutsch, Kimmo Eriksson, Dominique Foata, Ira Gessel, Curtis Greene, Patricia Hersh, Martin Isaacs, Benjamin Joseph, Martin Klazar, Donald Knuth, Darla Kremer, Valery Liskovets, Peter Littelmann, Ian Macdonald, Alexander Mednykh, Thomas Müller, Andrew Odlyzko, Alexander Postnikov, Robert Proctor, Douglas Rogers, Lou Shapiro, Rodica Simion, Mark Skandera, Louis Solomon, Dennis Stanton, Robert Sulanke, Sheila Sundaram, Jean-Yves Thibon, and Andrei Zelevinsky.

Richard Stanley
Cambridge, Massachusetts
March 1998

The paperback printing contains addenda to some of the exercises in a new section on page 583. The exercises in question are indicated by * in the main text.

---

p. 124, Exercise 5.28
p. 136, Exercise 5.41(j)
p. 144, Exercise 5.53
p. 151, Exercise 5.62(b)
p. 231, Exercise 6.25(i)
p. 232, Exercise 6.27(c)
p. 264, Exercise 6.19(iii)
p. 265, Exercise 6.19(mmm)
p. 272, Exercise 6.33(c)
p. 279, Exercise 6.56(c)
p. 467, Exercise 7.55(b)
p. 534, Exercise 7.74
p. 539, Exercise 7.85

Contents

Notation

The notation follows that of Volume 1, with the following exceptions.

- The coefficient of x^n in the power series $F(x)$ is now denoted $[x^n]F(x)$. This notation is generalized in an obvious way to such situations as

$$[x^m y^n]\sum_{i,j} a_{ij}x^i y^j = a_{mn}$$

$$\left[\frac{x^n}{n!}\right]\sum_i a_i \frac{x^i}{i!} = a_n.$$

- The number of inversions, number of descents, and major index of a permutation (or more generally of a sequence) w are denoted $\mathrm{inv}(w)$, $\mathrm{des}(w)$, and $\mathrm{maj}(w)$, respectively, rather than $i(w)$, $d(w)$, and $\iota(w)$. Sometimes, especially when we are regarding the symmetric group $\mathfrak{S}_n$ as a Coxeter group, we write $\ell(w)$ instead of $\mathrm{inv}(w)$.

The following notation is used for various rings and fields of generating functions. Here K denotes a field, which is always the field of coefficients of the series below. All Laurent series and fractional Laurent series are understood to have only finitely many terms with negative exponents.

$K[x]$	ring of polynomials in x
$K(x)$	field of rational functions in x (the quotient field of $K[x]$)
$K[[x]]$	ring of formal (power) series in x
$K((x))$	field of Laurent series in x (the quotient field of $K[[x]]$)
$K_{\mathrm{alg}}[[x]]$	ring of algebraic power series in x over $K(x)$
$K_{\mathrm{alg}}((x))$	field of algebraic Laurent series in x over $K(x)$
$K^{\mathrm{fra}}[[x]]$	ring of fractional power series in x

$K^{\text{fra}}((x))$	field of fractional Laurent series in x (the quotient field of $K^{\text{fra}}[[x]]$)
$K\langle X\rangle$	ring of noncommutative polynomials in the alphabet (set of variables) X
$K_{\text{rat}}\langle\langle X\rangle\rangle$	ring of rational (= recognizable) noncommutative series in the alphabet X
$K\langle\langle X\rangle\rangle$	ring of formal (noncommutative) series in the alphabet X
$K_{\text{alg}}\langle\langle X\rangle\rangle$	ring of (noncommutative) algebraic series in the alphabet X

5

Trees and the Composition of Generating Functions

5.1 The Exponential Formula

If $F(x)$ and $G(x)$ are formal power series with $G(0) = 0$, then we have seen (after Proposition 1.1.9) that the composition $F(G(x))$ is a well-defined formal power series. In this chapter we will investigate the combinatorial ramifications of power series composition. In this section we will be concerned with the case where $F(x)$ and $G(x)$ are exponential generating functions, and especially the case $F(x) = e^x$.

Let us first consider the combinatorial significance of the product $F(x)G(x)$ of two exponential generating functions

$$F(x) = \sum_{n\geq 0} f(n)\frac{x^n}{n!}$$

$$G(x) = \sum_{n\geq 0} g(n)\frac{x^n}{n!}.$$

Throughout this chapter K denotes a field of characteristic 0 (such as $\mathbb{C}$ with some indeterminates adjoined). We also denote by $E_f(x)$ the exponential generating function of the function $f : \mathbb{N} \to K$, i.e.,

$$E_f(x) = \sum_{n\geq 0} f(n)\frac{x^n}{n!}.$$

5.1.1 Proposition. *Given functions $f, g : \mathbb{N} \to K$, define a new function $h : \mathbb{N} \to K$ by the rule*

$$h(\#X) = \sum_{(S,T)} f(\#S)g(\#T), \tag{5.1}$$

where X is a finite set, and where (S, T) ranges over all weak ordered partitions of X into two blocks, i.e., $S \cap T = \emptyset$ and $S \cup T = X$. Then

$$E_h(x) = E_f(x)E_g(x). \tag{5.2}$$

Proof. Let $\#X = n$. There are $\binom{n}{k}$ pairs (S, T) with $\#S = k$ and $\#T = n - k$, so

$$h(n) = \sum_{k=0}^{n} \binom{n}{k} f(k) g(n-k).$$

From this (5.2) follows. □

One could also prove Proposition 5.1.1 by using Theorem 3.15.4 applied to the binomial poset $\mathbb{B}$ of Example 3.15.3.

We have stated Proposition 5.1.1 in terms of a certain relationship (5.1) among functions f, g, and h, but it is important to understand its combinatorial significance. Suppose we have two types of structures, say α and β, which can be put on a finite set X. We assume that the allowed structures depend only on the cardinality of X. A new "combined" type of structure, denoted $\alpha \cup \beta$, can be put on X by placing structures of type α and β on subsets S and T, respectively, of X such that $S \cup T = X$, $S \cap T = \emptyset$. If $f(k)$ (respectively $g(k)$) are the number of possible structures on a k-set of type α (respectively, β), then the right-hand side of (5.1) counts the number of structures of type $\alpha \cup \beta$ on X. More generally, we can assign a weight $w(\Gamma)$ to any structure Γ of type α or β. A combined structure of type $\alpha \cup \beta$ is defined to have weight equal to the product of the weights of each part. If $f(k)$ and $g(k)$ denote the sums of the weights of all structures on a k-set of types α and β, respectively, then the right-hand side of (5.1) counts the sum of the weights of all structures of type $\alpha \cup \beta$ on X.

5.1.2 Example. Given an n-element set X, let $h(n)$ be the number of ways to split X into two subsets S and T with $S \cup T = X$, $S \cap T = \emptyset$, and then to linearly order the elements of S and to choose a subset of T. There are $f(k) = k!$ ways to linearly order a k-element set, and $g(k) = 2^k$ ways to choose a subset of a k-element set. Hence

$$\begin{aligned} \sum_{n\geq 0} h(n)\frac{x^n}{n!} &= \left(\sum_{n\geq 0} n!\frac{x^n}{n!}\right)\left(\sum_{n\geq 0} 2^n\frac{x^n}{n!}\right) \\ &= \frac{e^{2x}}{1-x}. \end{aligned}$$

Proposition 5.1.1 can be iterated to yield the following result.

5.1.3 Proposition. *Fix $k \in \mathbb{P}$ and functions $f_1, f_2, \ldots, f_k : \mathbb{N} \to K$. Define a new function $h : \mathbb{N} \to K$ by*

$$h(\#S) = \sum f_1(\#T_1) f_2(\#T_2) \cdots f_k(\#T_k),$$

where $(T_1, \ldots, T_k)$ ranges over all weak ordered partitions of S into k blocks, i.e., $T_1, \ldots, T_k$ are subsets of S satisfying: (i) $T_i \cap T_j = \emptyset$ if $i \neq j$, and (ii) $T_1 \cup \cdots \cup T_k = S$. Then

$$E_h(x) = \prod_{i=1}^{k} E_{f_i}(x).$$

We are now able to give the main result of this section, which explains the combinatorial significance of the composition of exponential generating functions.

5.1.4 Theorem (The Compositional Formula). *Given functions $f : \mathbb{P} \to K$ and $g : \mathbb{N} \to K$ with $g(0) = 1$, define a new function $h : \mathbb{N} \to K$ by*

$$h(\#S) = \sum_{\pi=\{B_1,\ldots,B_k\}\in\Pi(S)} f(\#B_1)f(\#B_2)\cdots f(\#B_k)g(k), \quad \#S > 0, \qquad (5.3)$$

$$h(0) = 1,$$

where the sum ranges over all partitions (as defined in Section 1.4) $\pi = \{B_1, \ldots, B_k\}$ of the finite set S. Then

$$E_h(x) = E_g(E_f(x)).$$

(Here $E_f(x) = \sum_{n\geq 1} f(n)x^n/n!$, since f is only defined on positive integers.)

Proof. Suppose $\#S = n$, and let $h_k(n)$ denote the right-hand side of (5.3) for *fixed* k. Since $B_1, \ldots, B_k$ are nonempty, they are all distinct, so there are $k!$ ways of linearly ordering them. Thus by Proposition 5.1.3,

$$E_{h_k}(x) = \frac{g(k)}{k!} E_f(x)^k. \qquad (5.4)$$

Summing (5.4) over all $k \geq 1$ yields the desired result. □

Theorem 5.1.4 has the following combinatorial significance. Many structures on a set, such as graphs or posets, may be regarded as disjoint unions of their connected components. In addition, some additional structure may be placed on the components themselves, e.g., the components could be linearly ordered. If there are $f(j)$ connected structures on a j-set and $g(k)$ ways to place an additional structure on k components, then $h(n)$ is the total number of structures on an n-set. There is an obvious generalization to weighted structures, such as was discussed after Proposition 5.1.1.

The following example should help to elucidate the combinatorial meaning of Theorem 5.1.4; more substantial applications are given in Section 5.2.

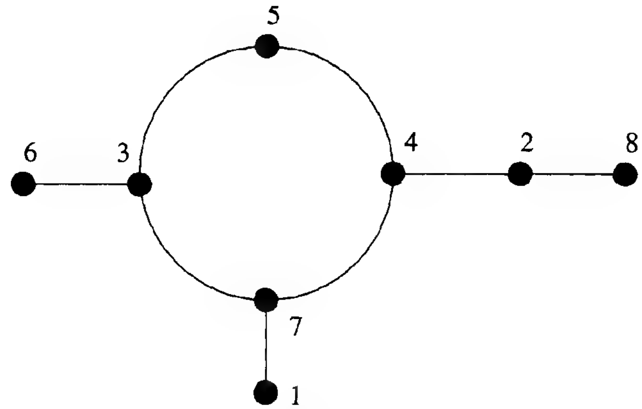


Figure 5-1. A circular arrangement of lines.

5.1.5 Example. Let $h(n)$ be the number of ways for n persons to form into nonempty lines, and then to arrange these lines in a circular order. Figure 5-1 shows one such arrangement of nine persons. There are $f(j) = j!$ ways to linearly order j persons, and $g(k) = (k-1)!$ ways to circularly order $k \geq 1$ lines. Thus

$$E_f(x) = \sum_{n\geq 1} n!\frac{x^n}{n!} = \frac{x}{1-x},$$

$$E_g(x) = 1 + \sum_{n\geq 1}(n-1)!\frac{x^n}{n!} = 1 + \log(1-x)^{-1},$$

so

$$\begin{aligned} E_h(x) &= E_g(E_f(x)) \\ &= 1 + \log\left(1 - \frac{x}{1-x}\right)^{-1} \\ &= 1 + \log(1-2x)^{-1} - \log(1-x)^{-1} \\ &= 1 + \sum_{n\geq 1}(2^n - 1)(n-1)!\frac{x^n}{n!}, \end{aligned}$$

whence $h(n) = (2^n - 1)(n-1)!$. Naturally, such a simple answer demands a simple combinatorial proof. Namely, arrange the n persons in a circle in $(n-1)!$ ways. In each of the n spaces between two persons, either do or do not draw a bar, except that at least one bar must be drawn. There are thus $2^n - 1$ choices for the bars. Between two consecutive bars (or a bar and itself if there is only one bar) read

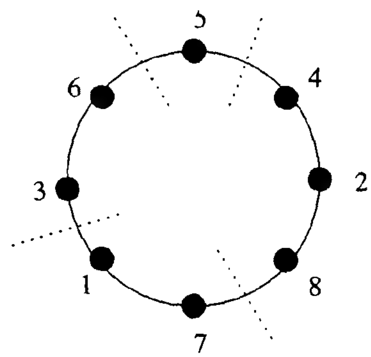


Figure 5-2. An equivalent form of Figure 5-1.

the persons in clockwise order to obtain their order in line. See Figure 5-2 for this method of representing Figure 5-1.

The most common use of Theorem 5.1.4 is the case where $g(k) = 1$ for all k. In combinatorial terms, a structure is put together from "connected" components, but no additional structure is placed on the components themselves.

5.1.6 Corollary (The Exponential Formula). *Given a function $f : \mathbb{P} \to K$, define a new function $h : \mathbb{N} \to K$ by*

$$\begin{aligned} h(\#S) &= \sum_{\pi=\{B_1,\ldots,B_k\}\in\Pi(S)} f(\#B_1)f(\#B_2)\cdots f(\#B_k), \qquad \#S > 0, &(5.5)\\ h(0) &= 1. \end{aligned}$$

Then

$$E_h(x) = \exp E_f(x). \tag{5.6}$$

Let us say a brief word about the computational aspects of equation (5.6). If the function $f(n)$ is given, then one can use (5.5) to compute $h(n)$. However, there is a much more efficient way to compute $h(n)$ from $f(n)$ (and conversely).

5.1.7 Proposition. *Let $f : \mathbb{P} \to K$ and $h : \mathbb{N} \to K$ be related by $E_h(x) = \exp E_f(x)$ (so in particular $h(0) = 1$). Then we have for $n \geq 0$ the recurrences*

$$h(n+1) = \sum_{k=0}^{n} \binom{n}{k} h(k) f(n+1-k), \tag{5.7}$$

$$f(n+1) = h(n+1) - \sum_{k=1}^{n} \binom{n}{k} h(k) f(n+1-k). \tag{5.8}$$

Proof. Differentiate $E_h(x) = \exp E_f(x)$ to obtain

$$E'_h(x) = E'_f(x)E_h(x). \tag{5.9}$$

Now equate coefficients of $x^n/n!$ on both sides of (5.9) to obtain (5.7). (It is also easy to give a combinatorial proof of (5.7).) Equation (5.8) is just a rearrangement of (5.7). □

The compositional and exponential formulas are concerned with structures on a set S obtained by choosing a partition of S and then imposing some "connected" structure on each block. In some situtations it is more natural to choose a *permutation* of S and then impose a "connected" structure on each cycle. These two situtations are clearly equivalent, since a permutation is nothing more than a partition with a cyclic ordering of each block. However, permutations arise often enough to warrant a separate statement. Recall that $\mathfrak{S}(S)$ denotes the set (or group) of all permutations of the set S.

5.1.8 Corollary (The Compositional Formula, permutation version). *Given functions $f : \mathbb{P} \to K$ and $g : \mathbb{N} \to K$ with $g(0) = 1$, define a new function $h : \mathbb{P} \to K$ by*

$$\begin{aligned} h(\#S) &= \sum_{\pi \in \mathfrak{S}(S)} f(\#C_1)f(\#C_2)\cdots f(\#C_k)g(k), \qquad \#S > 0, \\ h(0) &= 1, \end{aligned} \tag{5.10}$$

where $C_1, C_2, \ldots, C_k$ are the cycles in the disjoint cycle decomposition of π. Then

$$E_h(x) = E_g\left(\sum_{n\geq 1} f(n)\frac{x^n}{n}\right).$$

Proof. Since there are $(j-1)!$ ways to cyclically order a j-set, equation (5.10) may be written

$$h(\#S) = \sum_{\pi=\{B_1,\ldots,B_k\}\in\Pi(S)} [(\#B_1 - 1)!f(\#B_1)]\cdots[(\#B_k - 1)!f(\#B_k)]g(k),$$

so by Theorem 5.1.4,

$$\begin{aligned} E_h(x) &= E_g\left(\sum_{n\geq 1}(n-1)!f(n)\frac{x^n}{n!}\right) \\ &= E_g\left(\sum_{n\geq 1} f(n)\frac{x^n}{n}\right). \end{aligned}$$

□

5.1.9 Corollary (The Exponential Formula, permutation version). *Given a function $f : \mathbb{P} \to K$, define a new function $h : \mathbb{N} \to K$ by*

$$h(\#S) = \sum_{\pi \in \mathfrak{S}(S)} f(\#C_1) f(\#C_2) \cdots f(\#C_k), \qquad \#S > 0,$$
$$h(0) = 1,$$

where the notation is the same as in Corollary 5.1.8. Then

$$E_h(x) = \exp \sum_{n \geq 1} f(n) \frac{x^n}{n}.$$

In Section 3.15 [see Example 3.15.3(b)] we related addition and multiplication of exponential generating functions to the incidence algebra of the lattice of finite subsets of $\mathbb{N}$. There is a similar relation between *composition* of exponential generating functions and the incidence algebra of the lattice Π_n of partitions of $[n]$ (or any n-set). More precisely, we need to consider simultaneously all Π_n for $n \in \mathbb{P}$. Recall from Section 3.10 that if $\sigma \leq \pi$ in Π_n, then we have a natural decomposition

$$[\sigma, \pi] \cong \Pi_1^{a_1} \times \Pi_2^{a_2} \times \cdots \times \Pi_n^{a_n}, \tag{5.11}$$

where $|\sigma| = \sum i a_i$ and $|\pi| = \sum a_i$. Let $\mathbf{\Pi} = (\Pi_1, \Pi_2, \ldots)$. For each $n \in \mathbb{P}$, let $f_n \in I(\Pi_n, K)$, the incidence algebra of Π_n. Suppose that the sequence $f = (f_1, f_2, \ldots)$ satisfies the following property: There is a function (also denoted f) $f : \mathbb{P} \to K$ such that if $\sigma \leq \pi$ in Π_n and $[\sigma, \pi]$ satisfies (5.11), then

$$f_n(\sigma, \pi) = f(1)^{a_1} f(2)^{a_2} \cdots f(n)^{a_n}.$$

We then call f a *multiplicative function* on $\mathbf{\Pi}$.

For instance, if ζ_n is the zeta function of Π_n, then $\zeta = (\zeta_1, \zeta_2, \ldots)$ is multiplicative with $\zeta(n) = 1$ for all $n \in \mathbb{P}$. If μ_n is the Möbius function of Π_n, then by Proposition 3.8.2 and equation (3.30) we see that $\mu = (\mu_1, \mu_2, \ldots)$ is multiplicative with $\mu(n) = (-1)^{n-1}(n-1)!$.

Let $f = (f_1, f_2, \ldots)$ and $g = (g_1, g_2, \ldots)$, where $f_n, g_n \in I(\Pi_n, K)$. We can define the *convolution* $fg = ((fg)_1, (fg)_2, \ldots)$ by

$$(fg)_n = f_n g_n \quad (\text{convolution in } I(\Pi_n, K)). \tag{5.12}$$

5.1.10 Lemma. *If f and g are multiplicative on $\mathbf{\Pi}$, then so is fg.*

Proof. Let P and Q be locally finite posets, and let $u \in I(P, K)$, $v \in I(Q, K)$. Define $u \times v \in I(P \times Q, K)$ by

$$u \times v((x, x'), (y, y')) = u(x, y) v(x', y').$$

Then a straightforward argument as in the proof of Proposition 3.8.2 shows that $(u \times v)(u' \times v') = uu' \times vv'$. Thus from (5.11) we have

$$\begin{aligned}(fg)_n(\sigma, \pi) &= f_1 g_1(\hat{0}, \hat{1})^{a_1} \cdots f_n g_n(\hat{0}, \hat{1})^{a_n} \\ &= fg(1)^{a_1} \cdots fg(n)^{a_n}.\end{aligned}$$

□

It follows from Lemma 5.1.10 that the set $M(\Pi) = M(\Pi, K)$ of multiplicative functions on Π forms a semigroup under convolution. In fact, $M(\Pi)$ is even a monoid (= semigroup with identity), since the identity function $\delta = (\delta_1, \delta_2, \ldots)$ is multiplicative with $\delta(n) = \delta_{1n}$. (Caveat: $M(\Pi)$ is *not* closed under addition.)

5.1.11 Theorem. *Define a map $\phi : M(\Pi) \to xK[[x]]$ (the monoid of power series with zero constant term under composition) by*

$$\phi(f) = E_f(x) = \sum_{n \geq 1} f(n) \frac{x^n}{n!}.$$

Then ϕ is an anti-isomorphism of monoids, i.e., ϕ is a bijection and

$$\phi(fg) = E_g(E_f(x)).$$

Proof. Clearly ϕ is a bijection. Since fg is multiplicative by Lemma 5.1.10, it suffices to show that

$$\sum_{n \geq 1} fg(n) \frac{x^n}{n!} = E_g(E_f(x)).$$

By definition of $fg(n)$, we have in $I(\Pi_n, K)$

$$\begin{aligned} fg(n) &= \sum_{\pi = \{B_1, \ldots, B_k\} \in \Pi_n} f_n(\hat{0}, \pi) g_n(\pi, \hat{1}) \\ &= \sum_{\pi} f(\#B_1) \cdots f(\#B_k) g(k). \end{aligned} \tag{5.13}$$

Since (5.13) agrees with (5.3), the proof follows from Theorem 5.1.4. □

The next result follows from Theorem 5.1.11 in the same way that Proposition 3.15.5 follows from Theorem 3.15.4 (using Proposition 5.4.1), so the proof is omitted. (A direct proof avoiding Theorem 5.1.11 can also be given.) If $f = (f_1, f_2, \ldots)$ where $f_n \in I(\Pi_n, K)$ and each f_n^{-1} exists in $I(\Pi_n, K)$, then we write $f^{-1} = (f_1^{-1}, f_2^{-1}, \ldots)$.

5.1.12 Proposition. *Suppose f is multiplicative and f^{-1} exists. Then f^{-1} is multiplicative.*

5.1.13 Example. Let $\zeta, \delta, \mu \in M(\Pi)$ have the same meanings as above, so $\zeta\mu = \mu\zeta = \delta$. Now

$$E_\zeta(x) = \sum_{n\geq 1} \frac{x^n}{n!} = e^x - 1,$$

$$E_\delta(x) = x,$$

so by Theorem 5.1.11

$$\begin{aligned} &[\exp E_\mu(x)] - 1 = x \\ \Rightarrow \quad & E_\mu(x) = \log(1+x) \\ &\qquad = \sum_{n\geq 1} (-1)^{n-1}(n-1)!\frac{x^n}{n!} \\ \Rightarrow \quad & \mu(n) = (-1)^{n-1}(n-1)!. \end{aligned}$$

Thus we have another derivation of the Möbius function of Π_n (equation (3.30)).

5.1.14 Example. Let $h(n)$ be the number of ways to partition the set $[n]$, and then partition each block into blocks of odd cardinality. We are asking for the number of chains $\hat{0} \leq \pi \leq \sigma \leq \hat{1}$ in Π_n such that all block sizes of π are odd. Define $f \in M(\Pi)$ by

$$f(n) = \begin{cases} 1, & n \text{ odd}, \\ 0, & n \text{ even}. \end{cases}$$

Then clearly $h = f\zeta^2$, so by Theorem 5.1.11,

$$\begin{aligned} E_h(x) &= E_\zeta(E_\zeta(E_f(x))) \\ &= \exp\left[\left(\exp \sum_{n\geq 0} \frac{x^{2n+1}}{(2n+1)!}\right) - 1\right] - 1 \\ &= \exp(e^{\sinh x} - 1) - 1. \end{aligned}$$

We have discussed in this section the combinatorial significance of multiplying and composing exponential generating functions. Three further operations are important to understand combinatorially: addition, multiplication by x (really a special case of arbitrary multiplication, but of special significance), and differentiation.

5.1.15 Proposition. *Let S be a finite set. Given functions $f, g : \mathbb{N} \to K$, define new functions h_1, h_2, h_3, and h_4 as follows:*

$$h_1(\#S) = f(\#S) + g(\#S) \tag{5.14}$$

$$h_2(\#S) = (\#S)f(\#T), \quad \text{where } \#T = \#S - 1 \tag{5.15}$$

$$h_3(\#S) = f(\#T), \quad \text{where } \#T = \#S + 1 \tag{5.16}$$

$$h_4(\#S) = (\#S)f(\#S). \tag{5.17}$$

Then

$$E_{h_1}(x) = E_f(x) + E_g(x) \tag{5.18}$$

$$E_{h_2}(x) = xE_f(x) \tag{5.19}$$

$$E_{h_3}(x) = E'_f(x) \tag{5.20}$$

$$E_{h_4}(x) = xE'_f(x). \tag{5.21}$$

Proof. Easy. □

Equation (5.14) corresponds to a choice of two structures to place on S, one enumerated by f and one by g. In equation (5.15), we "root" a vertex v of S (i.e., we choose a distinguished vertex v, often called the *root*) and then place a structure on the remaining vertices $T = S - \{x\}$. Equation (5.16) corresponds to adjoining an extra element to S and then placing a structure enumerated by f. Finally in equation (5.17) we are simply placing a structure on S and rooting a vertex.

As we will see in subsequent sections, many structures have a recursive nature by which we can obtain from the results of this section functional equations or differential equations for the corresponding exponential generating function. Let us illustrate these ideas here by interpreting combinatorially the formula $E'_h(x) = E'_f(x)E_h(x)$ of equation (5.9). The left-hand side corresponds to the following construction: take a (finite) set S, adjoin a new element t, and then place on $S \cup \{t\}$ a structure enumerated by h (or *h-structure*). The right-hand side says: choose a subset T of S, adjoin an element t to T, place on $T \cup \{t\}$ an f-structure, and place on $S - T$ an h-structure. Clearly, if h and f are related by (5.5) (so that h-structures are unique disjoint unions of f-structures), then the combinatorial interpretations of $E'_h(x)$ and $E'_f(x)E_h(x)$ are equivalent.

5.2 Applications of the Exponential Formula

The most straightforward applications of Corollary 5.1.6 concern structures which have an obvious decomposition into "connected components."

5.2.1 Example. The number of graphs (without loops or multiple edges) on an n-element vertex set S is clearly $2^{\binom{n}{2}}$. (Each of the $\binom{n}{2}$ pairs of vertices may or may not be joined by an edge.) Let $c(\#S) = c(n)$ be the number of *connected* graphs on the vertex set S. Since a graph on S is obtained by choosing a partition π of S and then placing a connected graph on each block of π, we see that equation

(5.5) holds for $h(n) = 2^{\binom{n}{2}}$ and $f(n) = c(n)$. Hence by Corollary 5.1.6,

$$\begin{aligned} E_h(x) &= \sum_{n\geq 0} 2^{\binom{n}{2}} \frac{x^n}{n!} \\ &= \exp E_c(n) \\ &= \exp \sum_{n\geq 1} c(n) \frac{x^n}{n!}. \end{aligned}$$

Equivalently,

$$\sum_{n\geq 1} c(n)\frac{x^n}{n!} = \log \sum_{n\geq 0} 2^{\binom{n}{2}} \frac{x^n}{n!}. \tag{5.22}$$

Note that the generating functions $E_h(x)$ and $E_c(x)$ have zero radius of convergence; nevertheless, they still have combinatorial meaning.

Of course there is nothing special about graphs in the above example. If, for instance, $h(n)$ is the number of posets (or digraphs, topologies, triangle-free graphs, etc.) on an n-set and $c(n)$ is the number of connected posets (digraphs, topologies, triangle-free graphs, etc.) on an n-set, then the fundamental relation $E_h(x) = \exp E_c(x)$ continues to hold. In some cases (such as graphs and digraphs) we have an explicit formula for $h(n)$, but this is an incidental "bonus".

5.2.2 Example. Suppose we are interested in not just the number of connected graphs on an n-element vertex set, but rather the number of such graphs with exactly k components. Let $c_k(n)$ denote this number, and define

$$F(x,t) = \sum_{n\geq 0}\sum_{k\geq 0} c_k(n) t^k \frac{x^n}{n!}. \tag{5.23}$$

There are two ways to obtain this generating function from Theorem 5.1.4 and Corollary 5.1.6. We can either set $f(n) = c(n)$ and $g(k) = t^k$ in (5.3), or set $f(n) = c(n)t$ in (5.5). In either case we have

$$h(n) = \sum_{k\geq 0} c_k(n) t^k.$$

Thus

$$\begin{aligned} F(x,t) &= \exp t \sum_{n\geq 1} c(n) \frac{x^n}{n!} \\ &= \left(\sum_{n\geq 0} 2^{\binom{n}{2}} \frac{x^n}{n!} \right)^t. \end{aligned}$$

Again the same reasoning works equally as well for posets, digraphs, topologies, etc. In general, if $E_h(x)$ is the exponential generating function for the total number of structures on an n-set (where of course each structure is a unique disjoint union of connected components), then $E_h(x)^t$ also keeps track of the number of components, as in (5.23). Equivalently, if $h(n)$ is the number of structures on an n-set and $c_k(n)$ the number with k components, then

$$\begin{aligned}\sum_{k\geq 0} t^k E_{c_k}(x) &= E_h(x)^t \\ &= \exp t E_{c_1}(x) \\ &= \sum_{k\geq 0} t^k \frac{E_{c_1}(x)^k}{k!},\end{aligned} \tag{5.24}$$

so

$$E_{c_k}(x) = \frac{1}{k!} E_{c_1}(x)^k = \frac{1}{k!}[\log E_h(x)]^k,$$

where we set $c_k(0) = \delta_{0k}$ and $h(0) = 1$. In particular, if $h(n) = n!$ (the number of permutations of an n-set) then $E_h(x) = (1-x)^{-1}$, while $c_k(n) = c(n,k)$, the number of permutations of an n-set with k cycles. In other words, $c(n,k)$ is a signless Stirling number of the first kind (see Section 1.3); and we get

$$\sum_{n\geq 0} c(n,k)\frac{x^n}{n!} = \frac{1}{k!}[\log(1-x)^{-1}]^k. \tag{5.25}$$

Let us give one further "direct" application, concocted for the elegance of the final answer.

5.2.3 Example. Suppose we have a room containing n children. The children gather into circles by holding hands, and one child stands in the center of each circle. A circle may consist of as few as one child (clasping his or her hands), but each circle must contain a child inside it. In how many ways can this be done? Let this number be $h(n)$. An allowed arrangement of children is obtained by choosing a partition of the children, choosing a child c from each block B to be in the center of the circle, and arranging the other children in the block B in a circle about c. If $\#B = i \geq 2$, then there are $i \cdot (i-2)!$ ways to do this, and no ways if $i = 1$. Hence (setting $h(0) = 1$),

$$\begin{aligned}E_h(x) &= \exp \sum_{i\geq 2} i \cdot (i-2)! \frac{x^i}{i!} \\ &= \exp x \sum_{i\geq 1} \frac{x^i}{i} \\ &= \exp x \log(1-x)^{-1} \\ &= (1-x)^{-x}.\end{aligned}$$

Similarly, if $c_k(n)$ denotes the number of arrangements of n children with exactly k circles, then

$$\sum_{n\geq 0}\sum_{k\geq 0} c_k(n)t^k\frac{x^n}{n!} = (1-x)^{-xt}.$$

We next consider some problems concerned with successively partitioning the blocks of a partition.

5.2.4 Example. Let $B(n) = B_1(n)$ denote the n-th Bell number, i.e., $B(n) = \#\Pi_n$ (Section 1.4). Setting each $f(i) = 1$ in (5.5), we obtain

$$E_B(x) = \sum_{n\geq 0} B(n)\frac{x^n}{n!} = \exp(e^x - 1).$$

[See equation (1.24f).] Now let $B_2(n)$ be the number of ways to partition an n-set S, and then partition each block. Equivalently, $B_2(n)$ is the number of chains $\hat{0} \leq \pi_1 \leq \pi_2 \leq \hat{1}$ in Π_n. Putting each $f(i) = B(i)$ in (5.5), or equivalently, using Theorem 5.1.11 to compute $\phi(\zeta^3)$, we obtain

$$\sum_{n\geq 0} B_2(n)\frac{x^n}{n!} = \exp[\exp(e^x - 1) - 1].$$

Continuing in this manner, if $B_k(n)$ denotes the number of chains $\hat{0} \leq \pi_1 \leq \pi_2 \leq \cdots \leq \pi_k \leq \hat{1}$ in Π_n, then

$$\sum_{n\geq 0} B_k(n)\frac{x^n}{n!} = 1 + e^{\langle k+1\rangle}(x),$$

where $e(x) = e^{\langle 1\rangle}(x) = e^x - 1$ and $e^{\langle k+1\rangle}(x) = e(e^{\langle k\rangle}(x))$.

5.2.5 Example. The preceding example was quite straightforward. Consider now the following variation. Begin with an n-set S, and for $n \geq 2$ partition S into at least two blocks. Then partition each non-singleton block into at least two blocks. Continue partitioning each non-singleton block into at least two blocks, until only singletons remain. Call such a procedure a *total partition* of S. A total partition can be represented in a natural way by an (unordered) tree, as illustrated in Figure 5-3 for $S = [9]$. Notice that only the endpoints (leaves) need to be labeled; the other labels are superfluous. Let $t(n)$ denote the number of total partitions of S (with $t(0) = 0$). Thus $t(1) = 1,\ t(2) = 1,\ t(3) = 4,\ t(4) = 26$.

Consider what happens when we choose a partition π of S and then a total partition of each block of π. If $|\pi| = 1$, then we have done the equivalent of choosing a total partition of S. On the other hand, partitioning S into at least two blocks and then choosing a total partition of each block is equivalent to choosing a

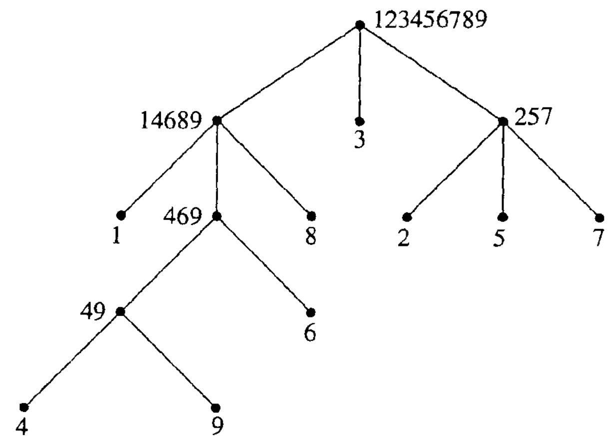


Figure 5-3. A total partition of [9] represented as a tree.

total partition of S itself. Thus altogether we obtain each total partition of S *twice*, provided $\#S \geq 2$. If $\#S = 1$, then we obtain the unique total partition of S only once. If $\#S = 0$ (i.e., $S = \emptyset$), then our procedure can be done in one way (i.e., do nothing), but by our convention there are no total partitions of S. Hence from Corollary 5.1.6 we obtain

$$\exp E_t(x) = 2E_t(x) - x + 1. \tag{5.26}$$

In other words, writing $F^{\langle -1\rangle}(x)$ for the compositional inverse of $F(x) = ax + bx^2 + \cdots$ where $a \neq 0$, i.e.,

$$F(F^{\langle -1\rangle}(x)) = F^{\langle -1\rangle}(F(x)) = x,$$

we have

$$E_t(x) = (1 + 2x - e^x)^{\langle -1\rangle}. \tag{5.27}$$

It does not seem possible to obtain a simpler result. In particular, in Section 5.4 we will discuss methods for computing the coefficients of compositional inverses, but these methods don't seem to yield anything interesting when $F(x) = 1 + 2x - e^x$. For some enumeration problems closely related to total partitions, see Exercises 5.26 and 5.40.

5.2.6 Example. Consider the variation of the preceding example where each non-singleton block must be partitioned into *exactly two* blocks. Call such a procedure a *binary total partition* of S, and denote the number of them by $b(\#S)$. As with total partitions, set $b(0) = 0$. The tree representing a binary total partition is a complete (unordered) binary tree, as illustrated in Figure 5-4. (Thus $b(n)$ is just the number

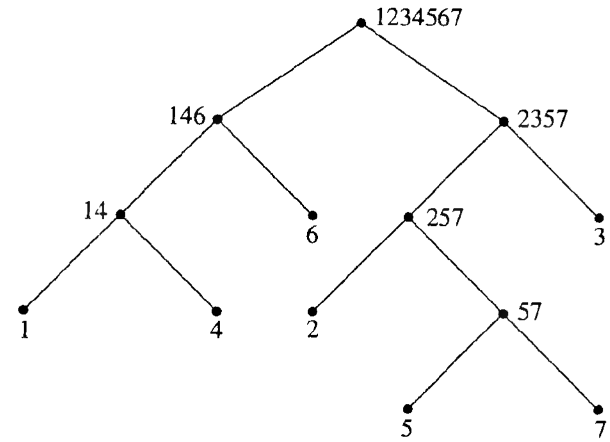


Figure 5-4. A binary total partition represented as a tree.

of (unordered) complete binary trees with n labeled endpoints.) It now follows from Theorem 5.1.4 (with $g(k) = \delta_{2k}$) or just by (5.4) (with $k = 2$ and $g(2) = 1$), in a similar way to how we obtained (5.26), that

$$\tfrac{1}{2}E_b(x)^2 = E_b(x) - x. \tag{5.28}$$

Solving this quadratic equation yields

$$\begin{aligned} E_b(x) &= 1 - \sqrt{1-2x} \\ &= 1 - \sum_{n\geq 0} \binom{\frac{1}{2}}{n} (-2)^n x^n \\ &= \sum_{n\geq 1} 1 \cdot 3 \cdot 5 \cdots (2n-3) \frac{x^n}{n!}, \end{aligned}$$

whence

$$b(n) = 1 \cdot 3 \cdot 5 \cdots (2n-3).$$

As usual, when such a simple answer is obtained, a direct combinatorial proof is desired. Now $1 \cdot 3 \cdot 5 \cdots (2n-3)$ is easily seen to be the number of partitions of $[2n-2]$ of type (2^{n-1}), i.e., with $n-1$ two-element blocks. Given a binary total partition β of $[n]$, we obtain a partition π of $[2n-2]$ of type (2^{n-1}) as follows. In the tree representing β (such as Figure 5-4), delete all the labels except the endpoints (leaves). Now iterate the following procedure until all vertices are labeled except the root. If labels $1, 2, \ldots, m$ have been used, then label by $m+1$ the vertex v satisfying: (a) v is unlabeled and both successors of v are labeled, and (b) among all unlabeled vertices with both successors labeled, the vertex having the successor

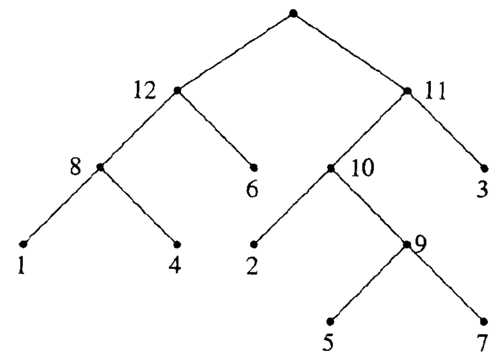


Figure 5-5. A decreasing labeled tree corresponding to a binary total partition.

with the *least* label is v. Figure 5-5 illustrates this procedure carried out for the tree in Figure 5-4. Finally let the blocks of π consist of the vertex labels of the two successors of a non-endpoint vertex. Thus from Figure 5-5 we obtain

$$\pi = \{\{1,4\}, \{2,9\}, \{3,10\}, \{5,7\}, \{6,8\}, \{11,12\}\}.$$

We leave it to the reader to check (not entirely trivial) that this procedure yields the desired bijection.

Certain problems involving symmetric matrices are well suited for use of the exponential formula. (Analogous results for arbitrary matrices are discussed in Section 5.5.) The basic idea is that a symmetric matrix $A = (a_{uv})$ whose rows and columns are indexed by a set V may be identified with a graph $G = G_A$ on the vertex set V, with an edge uv connecting u and v labeled by a_{uv}. (If $a_{uv} = 0$, then we simply omit the edge uv, rather than labeling it by 0. More generally, if $a_{uv} \in \mathbb{P}$, then it is often convenient to draw a_{uv} (unlabeled) edges between u and v.) Sometimes the connected components of G_A have a simple structure, so that the exponential formula can be used to enumerate all the graphs (or matrices).

5.2.7 Example. As in Proposition 4.6.21, let $S_n(2)$ denote the number of $n \times n$ symmetric $\mathbb{N}$-matrices A with every row (and hence every column) sum equal to two. The graph G_A has every vertex of degree two (counting loops once only). Hence the connected components of G_A must be (a) a single vertex with two loops, (b) a double edge between two vertices, (c) a cycle of length ≥ 3, or (d) a path of length ≥ 1 with a loop at each end. There are $\frac{1}{2}(n-1)!$ (undirected) cycles on $n \geq 3$ vertices, and $\frac{1}{2}n!$ (undirected) paths on $n \geq 2$ vertices with a loop at each end.

Hence by Corollary 5.1.6,

$$\begin{aligned}\sum_{n\geq 0} S_n(2)\frac{x^n}{n!} &= \exp\left(x + \frac{x^2}{2!} + \frac{1}{2}\sum_{n\geq 3}(n-1)!\frac{x^n}{n!} + \frac{1}{2}\sum_{n\geq 2} n!\frac{x^n}{n!}\right)\\ &= \exp\left(\frac{x^2}{4} + \frac{1}{2}\sum_{n\geq 1}\frac{x^n}{n} + \frac{1}{2}\sum_{n\geq 1}x^n\right)\\ &= \exp\left(\frac{x^2}{4} + \frac{1}{2}\log(1-x)^{-1} + \frac{x}{2(1-x)}\right)\\ &= (1-x)^{-1/2}\exp\left(\frac{x^2}{4} + \frac{x}{2(1-x)}\right).\end{aligned}$$

Using the technique of Exercise 5.24(c), we obtain the recurrence (writing $S_m = S_m(2)$)

$$S_{n+1} = (2n+1)S_n - (n)_2\, S_{n-1} - (n)_2\, S_{n-2} + \tfrac{1}{2}(n)_3\, S_{n-3}, \qquad n \geq 0.$$

5.2.8 Example. Suppose that in the previous example A must be a 0–1 matrix (i.e., the entry 2 is not allowed). Now the components of G_A of type (a) or (b) above are not allowed. If we let $S_n^*(2)$ denote the number of such matrices, it follows that

$$\begin{aligned}\sum_{n\geq 0} S_n^*(2)\frac{x^n}{n!} &= e^{-x-\frac{x^2}{2}}\sum_{n\geq 0} S_n(2)\frac{x^n}{n!}\\ &= (1-x)^{-1/2}\exp\left(-x - \frac{x^2}{4} + \frac{x}{2(1-x)}\right).\end{aligned}$$

As a further variation, suppose we again allow 2 as an entry, but that $\operatorname{tr} A = 0$ (i.e., all main-diagonal entries are zero). Now the components of G_A cannot have loops, so are of types (b) or (c). Hence, letting $T_n(2)$ be the number of such matrices, we have

$$\begin{aligned}\sum_{n\geq 0} T_n(2)\frac{x^n}{n!} &= \exp\left(\frac{x^2}{2!} + \frac{1}{2}\sum_{n\geq 3}(n-1)!\frac{x^n}{n!}\right)\\ &= (1-x)^{-1/2}\exp\left(-\frac{x}{2} + \frac{x^2}{4}\right).\end{aligned}$$

Similarly, if $T_n^*(2)$ denotes the number of traceless symmetric $n\times n$ 0–1 matrices with line sum 2, then

$$\begin{aligned}\sum_{n\geq 0} T_n^*(2)\frac{x^n}{n!} &= \exp\left(\frac{1}{2}\sum_{n\geq 3}(n-1)!\frac{x^n}{n!}\right)\\ &= (1-x)^{-1/2}\exp\left(-\frac{x}{2} - \frac{x^2}{4}\right). \qquad (5.29)\end{aligned}$$

The recurrence relations for $S_n^*(2)$, $T_n(2)$, and $T_n^*(2)$ turn out to be (using the technique of Exercise 5.24(c))

$$S_{n+1}^*(2) = 2n S_n^*(2) - (n)_2\, S_{n-1}^*(2) - \tfrac{1}{2}(n)_3\, S_{n-3}^*(2),$$

$$T_{n+1}(2) = nT_n(2) + nT_{n-1}(2) - \tbinom{n}{2}T_{n-2}(2),$$

$$T_{n+1}^*(2) = nT_n^*(2) + \tbinom{n}{2}T_{n-2}^*(2),$$

all valid for $n \geq 0$.

The next example is an interesting variation of the preceding two, where it is not *a priori* evident that the exponential formula is relevant.

5.2.9 Example. Let $X_n = (x_{ij})$ be an $n \times n$ symmetric matrix whose entries x_{ij} are independent indeterminates (over $\mathbb{R}$, say), except that $x_{ij} = x_{ji}$. Let $L(n)$ be the number of terms (i.e., distinct monomials) in the expansion of $\det X_n$, where we use only the variables x_{ij} for $i \leq j$. For instance,

$$\det X_3 = x_{11}x_{22}x_{33} + 2x_{12}x_{23}x_{13} - x_{13}^2x_{22} - x_{11}x_{23}^2 - x_{12}^2x_{33}.$$

Hence $L(3) = 5$, since the above sum has five terms. One might ask whether we should count a monomial that does arise in the expansion of $\det X_n$ but whose coefficient because of cancellation turns out to be zero. But we will soon see that no cancellation is possible; all occurrences of a given monomial have the same sign. Suppose now that $\pi \in \mathfrak{S}_n$. Define

$$M_\pi = x_{1,\pi(1)}x_{2,\pi(2)} \cdots x_{n,\pi(n)},$$

where we set $x_{ji} = x_{ij}$ if $j > i$. Thus M_π is the monomial corresponding to π in the expansion of $\det X_n$. Define a graph G_π whose vertex set is $[n]$, and with an (undirected) edge between i and $\pi(i)$ for $1 \leq i \leq n$. Thus the components of G_π are cycles of length ≥ 1. (A cycle of length 1 is a loop, and of length 2 is a double edge.) The crucial observation, whose easy proof we omit, is that $M_\pi = M_\sigma$ if and only if $G_\pi = G_\sigma$. Since a permutation π is even (respectively, odd) if and only if G_π has an even number (respectively, odd number) of cycles of even length, it follows that if $M_\pi = M_\sigma$ then M_π and M_σ occur in the expansion of $\det X_n$ with the same sign. In other words, there is no cancellation in the expansion of $\det X_n$. Also note that a graph G on $[n]$, every component of which is a cycle, is equal to G_π for some $\pi \in \mathfrak{S}_n$. (In fact, the number of such π is exactly $2^{d(\pi)}$, where π has $d(\pi)$ cycles of length ≥ 3.) It follows that $L(n)$ is simply the number of graphs on $[n]$ for which every component is a cycle (including loops and double edges).

Hence

$$\sum_{n\geq 0} L(n)\frac{x^n}{n!} = \exp\left(x + \frac{x^2}{2!} + \frac{1}{2}\sum_{n\geq 3}(n-1)!\frac{x^n}{n!}\right)$$
$$= (1-x)^{-1/2}\exp\left(\frac{x}{2} + \frac{x^2}{4}\right).$$

Note also that if P_π is the permutation matrix corresponding to $\pi \in \mathfrak{S}_n$, then $G_\pi = G_\sigma$ if and only if $P_\pi + P_\pi^{-1} = P_\sigma + P_\sigma^{-1}$. Hence $L(n)$ is the number of distinct matrices of the form $P_\pi + P_\pi^{-1}$, where $\pi \in \mathfrak{S}_n$. Equivalently, if we define $\pi, \sigma \in \mathfrak{S}_n$ to be *equivalent* if every cycle of π is a cycle or inverse of a cycle of σ, then $L(n)$ is the number of equivalence classes.

We conclude this section with some examples where it is more natural to use the permutation version of the exponential formula (Corollary 5.1.9).

5.2.10 Example. Let $\pi \in \mathfrak{S}_n$ be a permutation. Suppose that π has $c_i = c_i(\pi)$ cycles of length i, so $\sum ic_i = n$. Form a monomial

$$Z(\pi) = Z(\pi, t) = t_1^{c_1} t_2^{c_2} \cdots t_n^{c_n}$$

in the variables $t = (t_1, t_2, \ldots)$. We call $Z(\pi)$ the *cycle index* (or *cycle indicator* or *cycle monomial*) of π. Define the *cycle index* or *cycle index polynomial* (or *cycle indicator*, etc.) $Z(\mathfrak{S}_n)$ (also denoted $Z_{\mathfrak{S}_n}(t)$, $P_{\mathfrak{S}_n}(t)$, $\mathrm{Cyc}(\mathfrak{S}_n, t)$, etc.) of $\mathfrak{S}_n$ by

$$Z(\mathfrak{S}_n) = Z(\mathfrak{S}_n, t) = \frac{1}{n!}\sum_{\pi\in\mathfrak{S}_n} Z(\pi).$$

(In Section 7.24 we will consider the cycle index of other permutation groups.) It is sometimes more convenient to work with the *augmented cycle index*

$$\tilde{Z}(\mathfrak{S}_n) = n!\, Z(\mathfrak{S}_n) = \sum_{\pi\in\mathfrak{S}_n} Z(\pi).$$

For instance

$$\begin{aligned}
\tilde{Z}(\mathfrak{S}_1) &= t_1\\
\tilde{Z}(\mathfrak{S}_2) &= t_1^2 + t_2\\
\tilde{Z}(\mathfrak{S}_3) &= t_1^3 + 3t_1t_2 + 2t_3\\
\tilde{Z}(\mathfrak{S}_4) &= t_1^4 + 6t_1^2t_2 + 8t_1t_3 + 3t_2^2 + 6t_4.
\end{aligned}$$

Clearly, if we define $f : \mathbb{P} \to K$ by $f(n) = t_n$, then

$$\tilde{Z}(\mathfrak{S}_n) = \sum_{\pi \in \mathfrak{S}_n} f(\#C_1) f(\#C_2) \cdots f(\#C_k),$$

where $C_1, C_2, \ldots, C_k$ are the cycles of π. Hence by Corollary 5.1.9,

$$\sum_{n \geq 0} \tilde{Z}(\mathfrak{S}_n) \frac{x^n}{n!} = \exp \sum_{i \geq 1} t_i \frac{x^i}{i}. \tag{5.30}$$

There are many interesting specializations of (5.30) related to enumerative properties of $\mathfrak{S}_n$. For instance, fix $r \in \mathbb{P}$ and let $e_r(n)$ be the number of $\pi \in \mathfrak{S}_n$ satisfying $\pi^r = \mathrm{id}$, where id denotes the identity element of $\mathfrak{S}_n$. A permutation π satisfies $\pi^r = \mathrm{id}$ if and only if $c_d(\pi) = 0$ whenever $d \nmid r$. Hence

$$e_r(n) = Z(\mathfrak{S}_n) \Big|_{t_d = \begin{cases} 1, & d \mid r \\ 0, & d \nmid r. \end{cases}}$$

It follows that

$$\begin{aligned} \sum_{n \geq 0} e_r(n) \frac{x^n}{n!} &= \exp\left(\sum_{d \geq 1} t_d \frac{x^d}{d} \right) \Big|_{t_d = \begin{cases} 1, & d \mid r \\ 0, & d \nmid r \end{cases}} \\ &= \exp\left(\sum_{d \mid r} \frac{x^d}{d} \right). \end{aligned} \tag{5.31}$$

In particular, the number $e_2(n)$ of involutions in $\mathfrak{S}_n$ satisfies

$$\sum_{n \geq 0} e_2(n) \frac{x^n}{n!} = \exp\left(x + \frac{x^2}{2} \right). \tag{5.32}$$

This is the same generating function encountered way back in equation (1.10). Now we are able to understand its combinatorial significance more clearly.

There is a surprising connection between (a) Corollary 5.1.9, (b) the relationship between linear and circular words obtained in Proposition 4.7.11, and (c) the bijection $\pi \mapsto \hat{\pi}$ discussed in Section 1.3 between permutations written as products of cycles and as words. Basically, such a connection arises from a formula of the type

$$\sum_{n \geq 0} n!\, f(n) \frac{x^n}{n!} = \exp \sum_{n \geq 1} (n-1)!\, g(n) \frac{x^n}{n!}$$

because $n!$ is the number of linear words (permutations) on $[n]$, while $(n-1)!$ is the number of circular words (cycles). The next example may be regarded as the archetype for this line of thought.

5.2.11 Example. Let

$$F(x) = \prod_k (1 - t_k x)^{-1}, \tag{5.33}$$

where k ranges over some index set, say $k \in \mathbb{P}$. Thus

$$\begin{aligned}\log F(x) &= \sum_k \log(1 - t_k x)^{-1} \\ &= \sum_k \sum_{n \geq 1} t_k^n \frac{x^n}{n} \\ &= \sum_{n \geq 1} p_n(t) \frac{x^n}{n},\end{aligned} \tag{5.34}$$

where $p_n(t) = \sum_k t_k^n$. On the other hand, it is clear from (5.33) that

$$F(x) = \sum_{n \geq 0} h_n(t) x^n, \tag{5.35}$$

where $h_n(t)$ is the sum of all monomials of degree n in $t = (t_1, t_2, \ldots)$, i.e.,

$$h_n(t) = \sum_{\substack{a_1 + a_2 + \cdots = n \\ a_i \in \mathbb{N}}} t_1^{a_1} t_2^{a_2} \cdots.$$

(In Chapter 7 we will analyze the symmetric functions $p_n(t)$ and $h_n(t)$, as well as many others, in much greater depth.) From (5.34) and (5.35) we conclude

$$\sum_{n \geq 0} n!\, h_n(t) \frac{x^n}{n!} = \exp \sum_{n \geq 1} p_n(t) \frac{x^n}{n}. \tag{5.36}$$

We wish to give a direct combinatorial proof. By Corollary 5.1.9, the right-hand side is the exponential generating function for the following structure: Choose a permutation $\pi \in \mathfrak{S}_n$, and weight each cycle C of π by a monomial $t_k^{\#C}$ for some k. Define the total weight of π to be the product of the weights of each cycle. For instance, the list of structures of weight $u^2 v$ (where $u = t_1$ and $v = t_2$, say) is given by

$$\begin{array}{cc} \begin{matrix}(1)&(2)&(3)\\ u&u&v\end{matrix} & \begin{matrix}(12)&(3)\\ uu&v\end{matrix} \\[2ex] \begin{matrix}(1)&(2)&(3)\\ u&v&u\end{matrix} & \begin{matrix}(13)&(2)\\ uu&v\end{matrix} \\[2ex] \begin{matrix}(1)&(2)&(3)\\ v&u&u\end{matrix} & \begin{matrix}(1)&(23)\\ v&uu\end{matrix} \end{array} \tag{5.37}$$

Moreover, the left-hand side of (5.36) is clearly the exponential generating function for pairs (π, t^a), where $\pi \in \mathfrak{S}_n$ and t^a is a monomial of degree n in t. Thus the structures of weight u^2v are given by

$$\begin{array}{ll} 123,\ u^2v & 231,\ u^2v \\ 132,\ u^2v & 312,\ u^2v \\ 213,\ u^2v & 321,\ u^2v. \end{array} \tag{5.38}$$

In both (5.37) and (5.38) there are six items.

In general, in order to prove (5.36) bijectively, we need to do the following. Given a monomial t^a of degree n, find a bijection $\phi : \mathcal{C}_a \to \mathfrak{S}_n$, where $\mathcal{C}_a$ is the set of all permutations π in $\mathfrak{S}_n$, with each cycle C weighted by $t_j^{\#C}$ for some $j = j(C)$, such that the total weight $\Pi_C t_{j(C)}^{\#C}$ is equal to t^a. To describe ϕ, first impose some linear ordering on the t_j's, say $t_1 < t_2 < \cdots$. For fixed j, take all the cycles C of π with weight $t_j^{\#C}$ and write their standard representation (in the sense of Proposition 1.3.1), i.e., the largest element of each cycle is written first in the cycle, and the cycles are written left to right in increasing order of their largest elements. Remove the parentheses from this standard representation, obtaining a word w_j. Finally set $\phi(\pi) = (w, t^a)$, where $w = w_1 w_2 \cdots$ (juxtaposition of words). For instance, suppose π is the weighted permutation

$$\pi = \begin{array}{ccccc} (19) & (82) & (3) & (547) & (6) \\ vv & uu & v & uuu & u \end{array}$$

where $u < v$. The cycles weighted by u's and v's, respectively, have standard form

$$\begin{array}{ll} (6)(754)(82) & (\text{weight } u^6) \\ (3)(91) & (\text{weight } v^3). \end{array}$$

Hence

$$w_1 = 675482, \qquad w_2 = 391,$$
$$\phi(\pi) = (675482391, u^6v^3).$$

It is easy to check that ϕ is a bijection. Given (w, t^a), the monomial t^a determines the words w_j with their weights $t_j^{|w_j|}$. Each word w_j then corresponds to a collection of cycles C (with weight $t_j^{|C|}$) using the inverse of the bijection $\pi \mapsto \hat{\pi}$ of Section 1.3.

A similar argument leads to a direct combinatorial proof of equation (4.39); see Exercise 5.21.

5.3 Enumeration of Trees

Trees have a recursive structure which makes them highly amenable to the methods of this chapter. We will develop in this section some basic properties of trees as

a prelude to the Lagrange inversion formula of the next section. Trees are also fascinating objects of study for their own sake, so we will cover some topics not strictly germane to the composition of generating functions.

For the basic definitions and terminology concerning trees, see the Appendix of Volume 1. We also define a *planted forest* (also called a *rooted forest* or *forest of rooted trees*) to be a graph for which every connected component is a (rooted) tree. We begin with an investigation of the total number $p_k(n)$ of planted forests with k components on the vertex set $[n]$. Note that $p_1(n)$ is just the number $r(n)$ of rooted trees on $[n]$. If $S \subseteq [n]$ and $\#S = k$, then define $p_S(n)$ to be the number of planted forests on $[n]$ with k components, whose set of roots is S. Thus $p_k(n) = \binom{n}{k} p_S(n)$, since clearly $p_S(n) = p_T(n)$ if $\#S = \#T$.

5.3.1 Proposition. *Let*

$$y = R(x) = \sum_{n\geq 1} r(n)\frac{x^n}{n!},$$

where $r(n)$ as above is the number of rooted trees on the vertex set $[n]$ (with $r(0) = 0$). Then $y = xe^y$, or equivalently (since $x = ye^{-y}$),

$$y = (xe^{-x})^{\langle -1\rangle}. \tag{5.39}$$

Moreover, for $k \in \mathbb{P}$ we have

$$\frac{1}{k!}y^k = \sum_{n\geq 1} p_k(n)\frac{x^n}{n!}. \tag{5.40}$$

Proof. By Corollary 5.1.6, e^y is the exponential generating function for planted forests on the vertex set $[n]$. By equation (5.19), xe^y is the exponential generating function for the following structure on $[n]$. Choose a root vertex i, and place a planted forest F on the remaining vertices $[n] - \{i\}$. But this structure is equivalent to a tree with root i, whose subtrees of the root are the components of F. (See Figure 5-6.) Thus xe^y is just the exponential generating function for trees, so $y = xe^y$. Equation (5.40) then follows from Proposition 5.1.3. □

In the functional equation $y = xe^y$ of Proposition 5.3.1, substitute xe^y for the occurrence of y on the right-hand side to obtain

$$y = xe^{xe^y}$$

Again making the same substitution yields

$$y = xe^{xe^{xe^y}}.$$

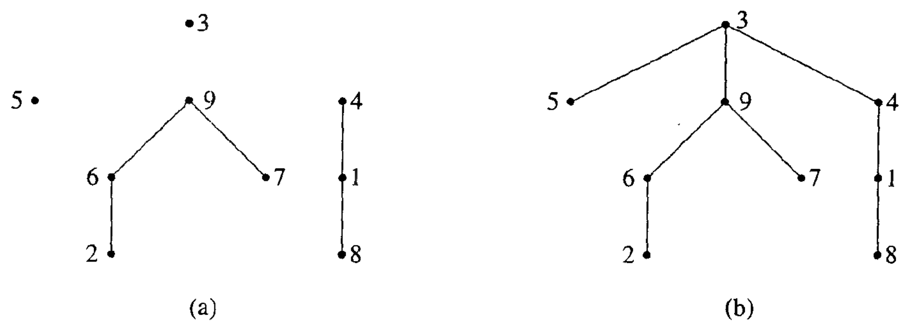


Figure 5-6. A rooted tree built from the subtrees of its root.

Iterating this procedure yields the "formula"

$$y = xe^{xe^{xe^{xe^{x^{\cdot^{\cdot^{\cdot}}}}}}} \quad . \tag{5.41}$$

The precise meaning of (5.41) is the following. Define $y_0 = x$ and for $k \geq 0$, $y_{k+1} = xe^{y_k}$. Then $\lim_{n\to\infty} y_k = y$, where the limit exists in the formal power series sense of Section 1.1. Moreover, by Corollary 5.1.6 and the second case of Proposition 5.1.15, we see that

$$y_k = \sum_{n\geq 1} r_k(n)\frac{x^n}{n!},$$

where $r_k(n)$ is the number of rooted trees on $[n]$ of length $\leq k$. For instance,

$$y_1 = xe^x = \sum_{n\geq 1} n\frac{x^n}{n!},$$

so $r_1(n) = n$ (as is obvious from the definition of $r_1(n)$).

The following quantities are closely related to the number $r(n)$ of rooted trees on the vertex set $[n]$:

$t(n)$ = number of free trees on $[n]$
$f(n)$ = number of free forests (i.e., disjoint unions of free trees) on $[n]$
$p(n)$ = number of planted forests on $[n]$.

We set $t(0) = 0$, $f(0) = 1$, $p(0) = 1$. Also write $T(x) = E_t(x)$, $F(x) = E_f(x)$,

and $P(x) = E_p(x)$. It is easy to verify the following relations:

$$\begin{aligned} r(n) &= np(n-1) = nt(n), \qquad p(n) = t(n+1), \\ F(x) &= e^{T(x)}, \qquad P(x) = e^{R(x)}, \\ P(x) &= T'(x), \qquad R(x) = xP(x). \end{aligned} \tag{5.42}$$

5.3.2 Proposition. *We have* $p_S(n) = kn^{n-k-1}$ *for any* $S \subseteq \binom{[n]}{k}$. *Thus*

$$\begin{aligned} p_k(n) &= k\binom{n}{k}n^{n-k-1} = \binom{n-1}{k-1}n^{n-k} \\ r(n) &= n^{n-1} \\ t(n) &= n^{n-2} \\ p(n) &= (n+1)^{n-1}. \end{aligned}$$

First Proof. The case $n = k$ is trivial, so assume $n \geq k+1$. The number of sequences $s = (s_1, \ldots, s_{n-k})$ with $s_i \in [n]$ for $1 \leq i \leq n-k-1$ and $s_{n-k} \in S$ is equal to kn^{n-k-1}. Hence we seek a bijection $\gamma : \mathcal{T}_{n,S} \to [n]^{n-k-1} \times S$, where $\mathcal{T}_{n,S}$ is the set of planted forests on $[n]$ with root set S. Given a forest $\sigma \in \mathcal{T}_{n,S}$, define a sequence $\sigma_1, \sigma_2, \ldots, \sigma_{n-k+1}$ of subforests (all with root set S) of σ as follows: Set $\sigma_1 = \sigma$. If $i < n-k+1$ and σ_i has been defined, then define σ_{i+1} to be the forest obtained from σ_i by removing its *largest nonroot endpoint* v_i (and the edge incident to v_i). Then define s_i to be the unique vertex of σ_i adjacent to v_i, and let $\gamma(\sigma) = (s_1, s_2, \ldots, s_{n-k})$. The sequence $\gamma(\sigma)$ is called the *Prüfer sequence* or *Prüfer code* of the planted forest σ. Figure 5-7 illustrates this construction with a forest $\sigma = \sigma_1 \in \mathcal{T}_{11,\{2,7\}}$ and the subforests σ_i, with vertex v_i circled. Hence for this example $\gamma(\sigma) = (5, 11, 5, 2, 9, 2, 7, 5, 7)$.

We claim that the map $\gamma : \mathcal{T}_{n,S} \to [n]^{n-k-1} \times S$ is a bijection. The crucial fact is that the largest element of $[n] - S$ missing from the sequence $(s_1, \ldots, s_{n-k})$ must be v_1 [why?]. Since v_1 and s_1 are adjacent, we are reduced to computing σ_2. But $\gamma(\sigma_2) = (s_2, s_3, \ldots, s_{n-k})$ (keeping in mind that the vertices of σ_2 are $[n] - \{v_1\}$,

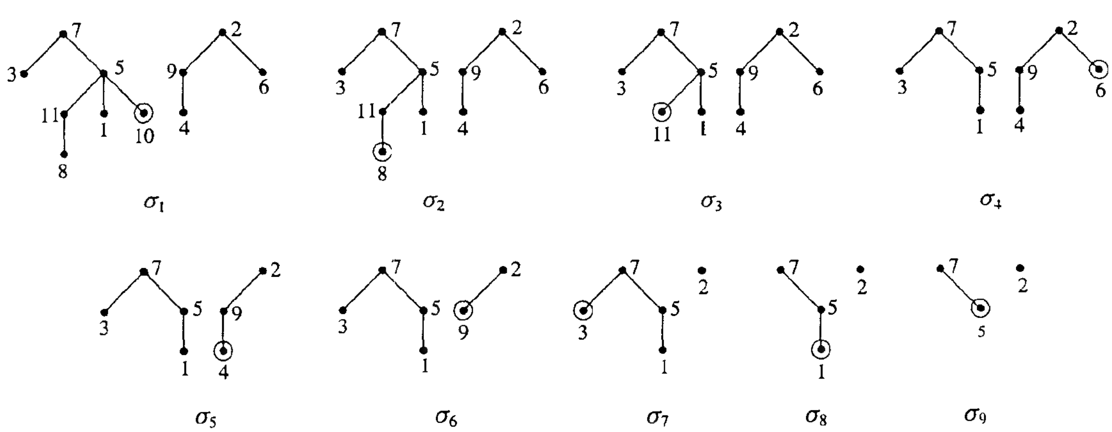


Figure 5-7. Constructing the Prüfer sequence of a labeled forest.

and not $[n-1]$). Hence by induction on n (the case $n = k+1$ being easy) we can recover σ uniquely from any $(s_1, \ldots, s_{n-k})$, so the proof is complete. □

5.3.3 Example. Let $S = \{2, 7\}$ and $(s_1, \ldots, s_9) = (5, 11, 5, 2, 9, 2, 7, 5, 7)$, so $n-k = 9$ and $n = 11$. The largest element of $[11]$ missing from $(s_1, \ldots, s_9)$ is 10. Hence 10 is an endpoint of σ adjacent to $s_1 = 5$. The largest element of $[11] - \{10\}$ missing from $(s_2, \ldots, s_9) = (11, 5, 2, 9, 2, 7, 5, 7)$ is 8. Hence 8 is an endpoint of σ_2 adjacent to $s_2 = 11$. The largest element of $[11] - \{8, 10\}$ missing from $(s_3, \ldots, s_9) = (5, 2, 9, 2, 7, 5, 7)$ is 11. Hence 11 is an endpoint of σ_3 adjacent to $s_3 = 5$. Continuing in this manner, we obtain the sequence of endpoints $10, 8, 11, 6, 4, 9, 3, 1, 5$. By beginning with the roots 2 and 7, and successively adding the endpoints in reverse order to the vertices $(s_9, \ldots, s_1) = (7, 5, 7, 2, 9, 2, 5, 11, 5)$, we obtain the forest $\sigma = \sigma_1$ of Figure 5-7.

Second proof of Proposition 5.3.2. We will show by a suitable bijection that

$$np_k(n) = k\binom{n}{k}n^{n-k}. \tag{5.43}$$

The underlying idea of the bijection is that a permutation can be represented both as a *word* and as a disjoint union of *cycles*. The bijection can be simplified somewhat for the case of rooted trees ($k = 1$), so we will present this special case first. Given a rooted tree τ on the vertex set $[n]$, circle a vertex $s \in [n]$. Let $w = w_1 w_2 \cdots w_k$ be the sequence (or word) of vertices in the unique path P in τ from the root r to s. Regard w as a permutation of its elements written in increasing order. For instance, if $w = 57283$, then w represents the permutation given by $w(2) = 5$, $w(3) = 7$, $w(5) = 2$, $w(7) = 8$, $w(8) = 3$, which in cycle notation is $(2, 5)(3, 7, 8)$. Let D_w be the directed graph with vertex set $A = \{w_1, \ldots, w_k\}$, and with an edge from j to $w(j)$ for all $j \in A$. Thus D_w is a disjoint union of (directed) cycles. When we remove from τ the edges of the path from r to s, we obtain a disjoint union of trees. Attach these trees to D_w by identifying vertices with the same label, and direct all the edges of these trees toward D_w. We obtain a digraph $D(\tau, s)$ for which all vertices have outdegree one. Moreover, the rooted tree τ, together with the distinguished vertex s, can be uniquely recovered from $D(\tau, s)$ by reversing the above steps. Since there are n choices for the vertex s, it follows that $nr(n)$ is equal to the number of digraphs on the vertex set $[n]$ for which every vertex has outdegree one. But such a digraph is just the digraph D_f of a function $f : [n] \to [n]$ (i.e., for each $j \in [n]$ draw an edge from j to $f(j)$). Since there are n^n such functions, we get $nr(n) = n^n$, so $r(n) = n^{n-1}$ as desired.

If we try the same idea for arbitrary planted forests σ, we end up needing to count functions f from some subset B of $[n]$ to $[n]$ such that the digraph D_f with vertex set $[n]$ and edges $j \to f(j)$ is *nonacyclic* (i.e., has at least one directed cycle). Since there is no obvious way to count such functions, some modification

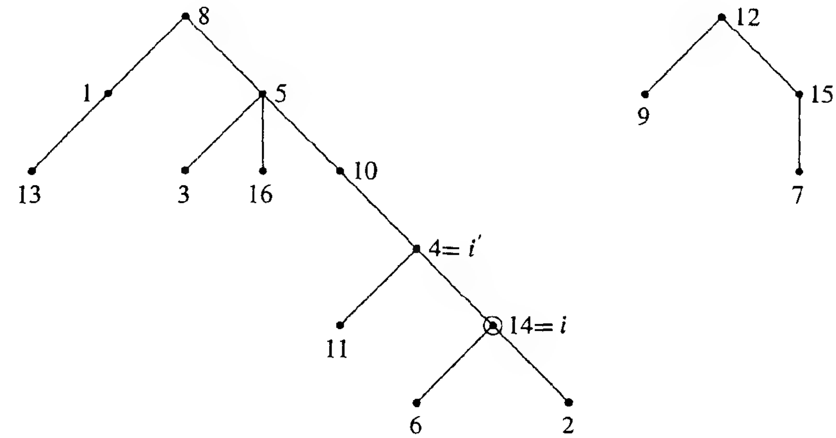


Figure 5-8. An illustration of the second proof of Proposition 5.3.2.

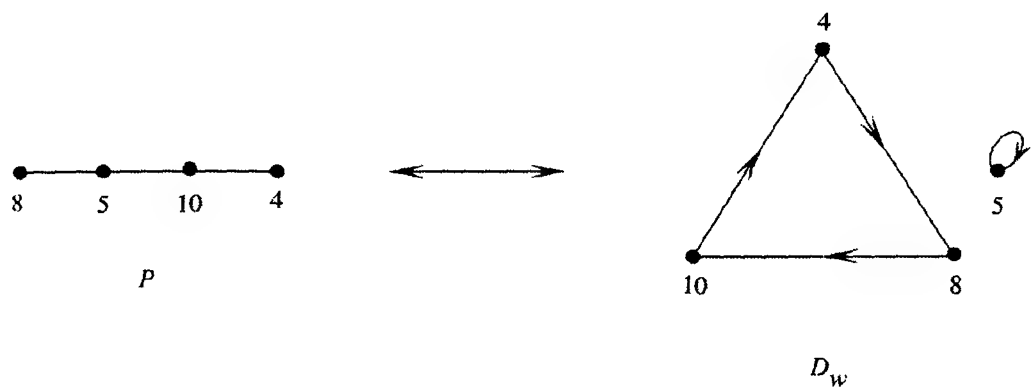


Figure 5-9. Two graphical representations of a permutation.

of the above bijection is needed. Note that the nonacyclicity condition is irrelevant when $k = 1$, since when $B = [n]$ the digraph D_f is always nonacyclic.

We now proceed to the correct bijection in the general case. Let σ be a planted forest on $[n]$ with k components. Circle a vertex i of σ. Figure 5-8 illustrates the case $n = 16$, $k = 2$, $i = 14$. The vertex i belongs to a component τ of σ. Remove from τ the complete subtree τ_i with root i (keeping i circled). If i is not the root of τ then let i' be the unique predecessor of i in τ. (If i is a root, then ignore all steps below involving i'.) Let $w = w_1 w_2 \cdots w_k$ be the sequence (or word) of vertices in the unique path P in $\tau - \tau_i$ from the root r to i'. Let $A = \{w_1, \ldots, w_k\}$ be the set of vertices of P. In the example of Figure 5-8, we have $w = 8, 5, 10, 4$. Regard w as a permutation of its elements written in increasing order. For our example, the permutation is given by $w(4) = 8$, $w(5) = 5$, $w(8) = 10$, $w(10) = 4$, which in cycle notation is $(4, 8, 10)(5)$. Let D_w be the directed graph with vertex set A, and with an edge from j to $w(j)$ for all $j \in A$. (See Figure 5-9.)

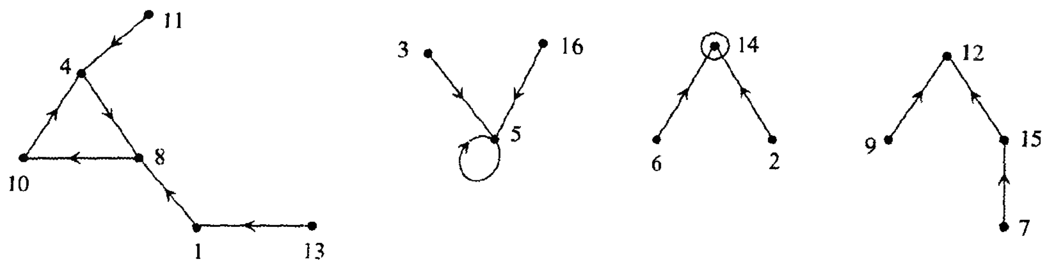


Figure 5-10. The digraph $D(\sigma, i)$.

When we remove from $\tau - \tau_i$ the edges of the path P, we obtain a collection of (rooted) trees whose roots are the vertices in P. Attach these trees to D_w by identifying vertices with the same label. Direct all the edges of these trees toward D_w. For each component of σ other than τ, and for τ_i, direct their edges toward the root. We obtain a digraph $D(\sigma, i)$ on $[n]$ for which k vertices have outdegree zero and the remaining $n-k$ vertices have outdegree one. Moreover, one of the vertices of outdegree zero is circled. (See Figure 5-10.) If i is a root of τ, then $D(\sigma, i)$ is just σ with all edges directed toward roots.

Let B be the set of vertices of $D(\sigma, i)$ of outdegree one. We may identify $D(\sigma, i)$ with the function $f : B \to [n]$ defined by $f(a) = b$ if $a \to b$ is an edge of $D(\sigma, i)$. Moreover, the circled vertex i belongs to $[n] - B$.

It is not difficult to reverse all the steps and obtain the pair (σ, i) from (f, i). There are $np_k(n)$ pairs (σ, i). We can choose B to be any $(n-k)$-subset of $[n]$ in $\binom{n}{k}$ ways, then choose $i \in [n] - B$ in k ways, and finally choose $f : B \to [n]$ in n^{n-k} ways. Hence (5.43) follows. □

The surprising formula

$$R(xe^{-x}) = \sum_{n\geq 1} n^{n-1} \frac{(xe^{-x})^n}{n!} = x, \tag{5.44}$$

inherent in equation (5.39) and the formula $r(n) = n^{n-1}$ of Proposition 5.3.2, can be proved directly as follows:

$$\begin{aligned}
\sum_{n\geq 1} n^{n-1} \frac{(xe^{-x})^n}{n!} &= \sum_{n\geq 1} \frac{n^{n-1}x^n}{n!} \sum_{k\geq 0} \frac{(-nx)^k}{k!} \\
&= \sum_{m\geq 1} \frac{x^m}{m!} \sum_{j=1}^{m} \binom{m}{j} (-1)^{m+j} j^{m-1} \\
&= \sum_{m\geq 1} \frac{x^m}{m!} [\Delta^m 0^{m-1} - (-1)^m 0^{m-1}],
\end{aligned} \tag{5.45}$$

by applying (1.27) to the function $f(j) = j^{m-1}$. Here we must interpret $0^0 = 1$. Then by Proposition 1.4.2(a), the sum in (5.45) collapses to the single term x.

The two proofs of Proposition 5.3.2 lead to an elegant refinement of the formula $r(n) = n^{n-1}$. Given a vertex v of a planted forest σ, define the *degree* $\deg v$ of v to be the number of successors of v. Thus v is an endpoint of σ if and only if $\deg v = 0$. If the vertex set of σ is $[n]$, then define the *ordered degree sequence* $\Delta(\sigma) = (\delta_1, \dots, \delta_n)$, where $\delta_i = \deg i$. It is easy to see that a sequence $(\delta_1, \dots, \delta_n) \in \mathbb{N}^n$ is the ordered degree sequence of some planted forest σ on $[n]$ with k components if and only if

$$\sum_{i=1}^{n} \delta_i = n - k. \tag{5.46}$$

5.3.4 Theorem. *Let $\delta = (\delta_1, \dots, \delta_n) \in \mathbb{N}^n$ with $\sum \delta_i = n - k$. The number $N(\delta)$ of planted forests σ on the vertex set $[n]$ (necessarily with k components) with ordered degree sequence $\Delta(\sigma) = \delta$ is given by*

$$N(\delta) = \binom{n-1}{k-1}\binom{n-k}{\delta_1, \delta_2, \dots, \delta_n}.$$

Equivalently,

$$\sum_{\sigma} x_1^{\deg 1} \cdots x_n^{\deg n} = \binom{n-1}{k-1}(x_1 + \cdots + x_n)^{n-k}, \tag{5.47}$$

where σ ranges over all planted forests on $[n]$ with k components.

First Proof. Consider the first proof of Proposition 5.3.2. The number of times $j \in [n]$ appears in the sequence $\gamma(\sigma)$ is clearly equal to $\deg j$, since j is the predecessor of exactly $\deg j$ vertices v_i. Hence for fixed root set S,

$$\sum_{\sigma \in T_{n,S}} x_1^{\deg 1} \cdots x_n^{\deg n} = (x_1 + \cdots + x_n)^{n-k-1} \sum_{i \in S} x_i.$$

Now sum over all $S \in \binom{[n]}{k}$ to obtain (5.47). $\square$

Second Proof. Now consider the second proof of Proposition 5.3.2. The key observation here is that for each $j \in [n]$, the degree of vertex j in the planted forest σ is equal to the indegree of j in the digraph $D(\sigma, i)$, or equivalently, $\deg j = \#f^{-1}(j)$. Hence

$$n \sum_{\sigma} x_1^{\deg 1} \cdots x_n^{\deg n} = k \sum_{\substack{B \subseteq [n] \\ \#B = n-k}} \sum_{f : B \to [n]} x_1^{\#f^{-1}(1)} \cdots x_n^{\#f^{-1}(n)}, \tag{5.48}$$

where σ ranges over all k-component planted forests on $[n]$. The inner sum in the right-hand side of (5.48) is independent of B and is equal to $(x_1 + \cdots + x_n)^{n-k}$.

Hence

$$n\sum_{\sigma} x_1^{\deg 1}\cdots x_n^{\deg n} = k\binom{n}{k}(x_1+\cdots+x_n)^{n-k},$$

which is equivalent to (5.47). □

There is an alternative way of stating Theorem 5.3.4 that is sometimes more convenient. Given a planted forest σ, define the *type* of σ to be the sequence

$$\text{type }\sigma = (r_0, r_1, \ldots),$$

where r_i vertices of σ have degree i. We also write type $\sigma = (r_0, r_1, \ldots, r_m)$ if $r_j = 0$ for $j > m$. It follows easily from (5.46) that a sequence $\mathbf{r} = (r_0, r_1, \ldots)$ of nonnegative integers is the type of some planted forest with n vertices and k components if and only if

$$\sum r_i = n, \qquad \sum (i-1)r_i = -k. \tag{5.49}$$

5.3.5 Corollary. *Let* $\mathbf{r} = (r_0, r_1, \ldots)$ *be a sequence of nonnegative integers satisfying (5.49). Then the number* $M(\mathbf{r})$ *of planted forests* σ *on the vertex set* $[n]$ *(necessarily with* k *components) of type* $\mathbf{r}$ *is given by*

$$\begin{aligned} M(\mathbf{r}) &= \binom{n-1}{k-1}\frac{(n-k)!}{0!^{r_0}1!^{r_1}\cdots}\binom{n}{r_0, r_1, \ldots} \\ &= \frac{k}{n}\binom{n}{k}\frac{(n-k)!}{0!^{r_0}1!^{r_1}\cdots}\binom{n}{r_0, r_1, \ldots}. \end{aligned}$$

We have been considering up to now the case of *labeled trees*, i.e., trees whose vertices are distinguishable. We next will deal with *unlabeled plane forests* σ, so the vertices of σ are regarded as indistinguishable, but the subtrees at any vertex (as well as the components themselves of σ) are linearly ordered. This automatically makes the vertices of σ distinguishable (in other words, an unlabeled plane forest has only the trivial automorphism), so it really makes no difference whether or not the vertices of σ are labeled. (An unlabeled plane forest with n vertices has $n!$ labelings.) Thus all plane forests will henceforth be assumed to be unlabeled. We continue to define the *degree* of a vertex v to be the number of successors (children) of v, and the *type* of σ is $\mathbf{r} = (r_0, r_1, \ldots)$ if r_i vertices have degree i. Equation (5.49) continues to be the condition on nonnegative integers $r_0, r_1, \ldots$ for there to exist a plane forest with n vertices, k components, and type $\mathbf{r} = (r_0, r_1, \ldots)$. Thus, for example, Figure 5-11 shows the ten plane trees of type (3, 1, 2), while Figure 5-12 illustrates a plane forest with 12 vertices, 3 components, and type (7, 2, 2, 1).

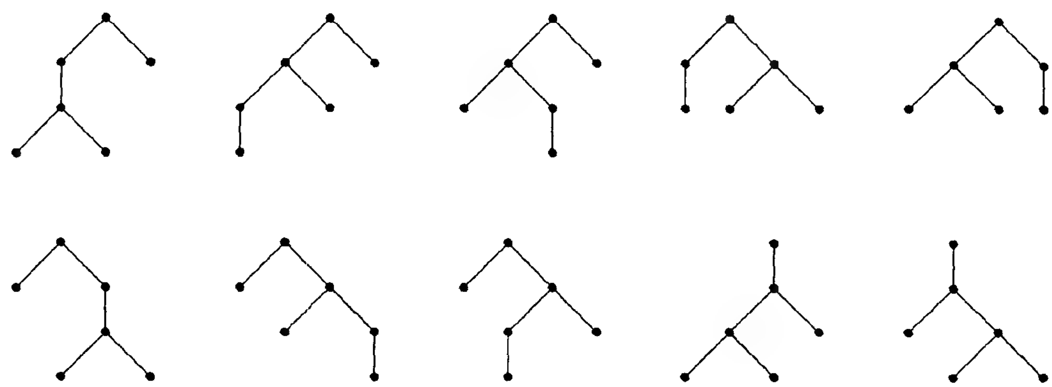


Figure 5-11. The ten plane trees of type (3, 1, 2).

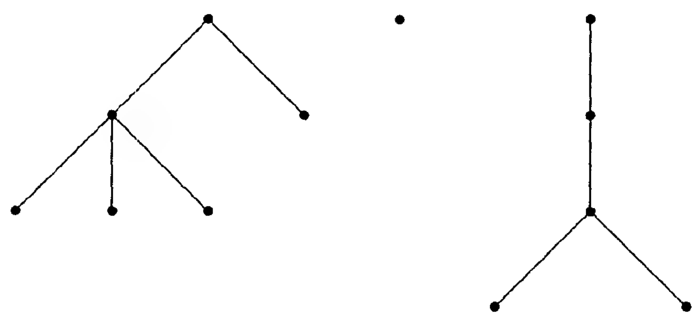


Figure 5-12. A plane forest of type (7, 2, 2, 1).

Our goal here is to enumerate unlabeled plane forests of a given type. This result will be used in the next section to prove the Lagrange inversion formula. It is convenient to work in the context of words in free monoids, as discussed in Section 4.7. Our alphabet $\mathcal{A}$ will consist of letters $x_0, x_1, x_2, \ldots$. (For plane forests with maximum degree m, it will suffice to take $\mathcal{A} = \{x_0, \ldots, x_m\}$.) The empty word is denoted by 1. Define the *weight* $\phi(x_i)$ of the letter x_i by $\phi(x_i) = i - 1$, and extend ϕ to $\mathcal{A}^*$ by

$$\phi(w_1 w_2 \cdots w_j) = \phi(w_1) + \phi(w_2) + \cdots + \phi(w_j),$$

where each $w_i \in \mathcal{A}$. (Set $\phi(1) = 0$.) Define a subset $\mathcal{B} \subset \mathcal{A}^*$ by

$$\mathcal{B} = \{w \in \mathcal{A}^* : \phi(w) = -1; \text{and if } w = uv \text{ where } v \neq 1, \text{then } \phi(u) \geq 0\}. \quad (5.50)$$

The elements of $\mathcal{B}$ are called *Łukasiewicz words*; see Example 6.6.7 for further information.

5.3.6 Lemma. *The monoid $\mathcal{B}^*$ generated by $\mathcal{B}$ is very pure (and hence free) with basis $\mathcal{B}$. (See Section 4.7 for relevant definitions.)*

Proof. Let $w = w_1 \cdots w_m \in \mathcal{B}^*$, where $w_i \in \mathcal{A}$. Let j be the least integer for which $\phi(w_1 \cdots w_j) < 0$, so in fact $\phi(w_1 \cdots w_j) = -1$ and $u = w_1 \cdots w_j \in \mathcal{B}$. Clearly if $w = vv'$ with $v \in \mathcal{B}$ then $u = v$. Thus by induction on the length of w, we obtain a unique factorization of w into elements of $\mathcal{B}$, so $\mathcal{B}^*$ is free with basis $\mathcal{B}$.

To show that $\mathcal{B}^*$ is very pure, it suffices to show [why?] that we cannot have $u, v, w \in \mathcal{A}^+ := \mathcal{A}^* - \{1\}$ with $uv \in \mathcal{B}$ and $vw \in \mathcal{B}$. But if $uv \in \mathcal{B}$ then $\phi(u) \geq 0$ and $\phi(u)+\phi(v) = -1$, so $\phi(v) < 0$. This contradicts $vw \in \mathcal{B}$, so $\mathcal{B}^*$ is very pure. □

Recall from Section 4.7 that if $w = w_1 w_2 \cdots w_m \in \mathcal{A}^*$ with $w_i \in \mathcal{A}$, then a cyclic shift $w_i w_{i+1} \cdots w_m w_1 \cdots w_{i-1}$ of w is called a *conjugate* (or $\mathcal{A}$-conjugate if there is a possibility of confusion) of w. (The reason for this terminology is that in a group G, the elements $w_1 w_2 \cdots w_m$ and $w_i w_{i+1} \cdots w_{i-1}$ are conjugate in the usual group-theoretic sense.)

5.3.7 Lemma. *A word $w \in \mathcal{A}^*$ is a conjugate of a word in $\mathcal{B}^*$ if and only if $\phi(w) < 0$.*

First Proof. Since $\phi(w)$ is unaffected by conjugation, clearly $\phi(w) < 0$ for every conjugate of a word in $\mathcal{B}^*$. We show the converse by induction on the length (in $\mathcal{A}^*$) $\ell(w)$ of w. The assertion is clear for $\ell(w) = 0$ (so $w = 1$), so assume it for words of length $< m$ and let $w = w_1 \cdots w_m$ where $w_i \in \mathcal{A}$ and $\phi(w) < 0$. Since $\phi(w_1) + \cdots + \phi(w_m) < 0$ and since $\phi(w_i) < 0$ only when $\phi(w_i) = -1$, it is easily seen that some conjugate w' of w has the form $w' = x_{s+1} x_0^s v$ for some $s \geq 0$. Since $\phi(v) = \phi(w') < 0$, it follows by induction that some conjugate v' of v lies in $\mathcal{B}^*$. Specifically, say that $v = yz$ where $zy \in \mathcal{B}^*$ and $y \neq 1$ (so that if v itself is in $\mathcal{B}^*$, then we take $y = v$ and $z = 1$). But then it is easily seen that $z x_{s+1} x_0^s y \in \mathcal{B}^*$. Since $z x_{s+1} x_0^s y$ is a conjugate of w, the proof follows by induction. □

Second Proof (sketch). The previous proof was straightforward but not particularly enlightening. We sketch another proof based on geometrical considerations which is more intuitive. Given any word $u = u_1 \cdots u_m \in \mathcal{A}^*$, with $u_i \in \mathcal{A}$, associate with u a lattice path *LP(u)* with m steps in $\mathbb{R}^2$ as follows. Begin at $(0, 0)$, and let the i-th step, for $1 \leq i \leq m$, be $(1, \phi(u_i))$. Now suppose $w \in \mathcal{A}^*$ and $\phi(w) < 0$, and consider the path $LP(w^2)$. Figure 5-13 illustrates $LP(w^2)$ for $w = x_0 x_1 x_0^2 x_2 x_0^2 x_2 x_0 x_3$. Suppose $\phi(w) = -k$. Let B be the leftmost lowest point on $LP(w^2)$, and let A be the leftmost point which is exactly k levels higher than B. (See Figure 5-13.) The horizontal distance between A and B is exactly m. If we translate the part of $LP(w^2)$ between A and B so that A is at the origin, then the resulting path is equal to $LP(v)$, where v is a conjugate of w belonging to $\mathcal{B}^*$. □

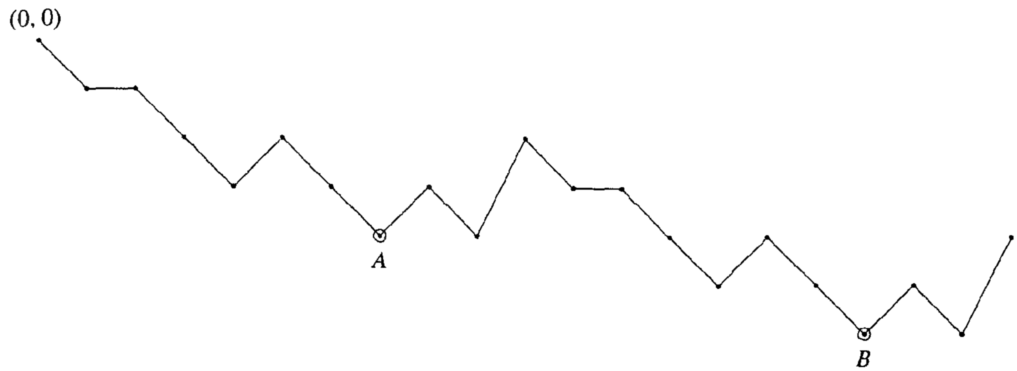


Figure 5-13. A lattice path $LP(w^2)$.

5.3.8 Example. Let $w = x_0x_1x_0^2x_2x_0^2x_2x_0x_3$ as in the second proof above. From Figure 5-13 we see that if A is translated to $(0, 0)$ then the path between A and B is $LP(v)$, where $v = x_2x_0x_3x_0x_1x_0^2x_2x_0^2$. The unique factorization of v into elements of $\mathcal{B}$ is

$$v = (x_2x_0x_3x_0x_1x_0^2)(x_2x_0^2).$$

Since $\phi(w) = -2$ and $\mathcal{B}^*$ is very pure, there are precisely two conjugates of w that belong to $\mathcal{B}^*$, viz., v and

$$u = (x_2x_0^2)(x_2x_0x_3x_0x_1x_0^2).$$

In general, if $\phi(w) = -k$ then precisely k conjugates of w belong to $\mathcal{B}^*$. However, these conjugates might not be all distinct elements of $\mathcal{A}^*$. For instance, if $w = x_0^k$ then all k conjugates of w are equal to w.

We now wish to associate with an unlabeled plane forest σ with n vertices a word $w(\sigma)$ in $\mathcal{A}^*$ of length n (and weight $\phi(w(\sigma)) = -k$, where σ has k components). To do this, we first need to define a certain canonical linear ordering on the vertices of σ, called *depth-first order* or *preorder*, and denoted $\mathrm{ord}(\sigma)$. It is defined recursively as follows:

(a) If σ has $k \geq 2$ components $\tau_1, \ldots, \tau_k$ (listed in the order defining σ as a plane forest), then set

$$\mathrm{ord}(\sigma) = \mathrm{ord}(\tau_1), \ \ldots, \ \mathrm{ord}(\tau_k) \qquad \text{(concatenation of words).}$$

(b) If σ has one component, then let $\tau_1, \ldots, \tau_j$ be the subtrees of the root v (listed in the order defining σ as a plane tree). Set

$$\mathrm{ord}(\sigma) = v, \ \mathrm{ord}(\tau_1), \ \ldots, \ \mathrm{ord}(\tau_k) \qquad \text{(concatenation of words).}$$

The preorder on a plane tree has an alternative informal description as follows. Imagine that the edges of the tree are wooden sticks, and that a worm begins just left of the root and crawls along the outside of the sticks until (s)he (or it) returns to the starting point. Then the order in which vertices are seen for the first time is

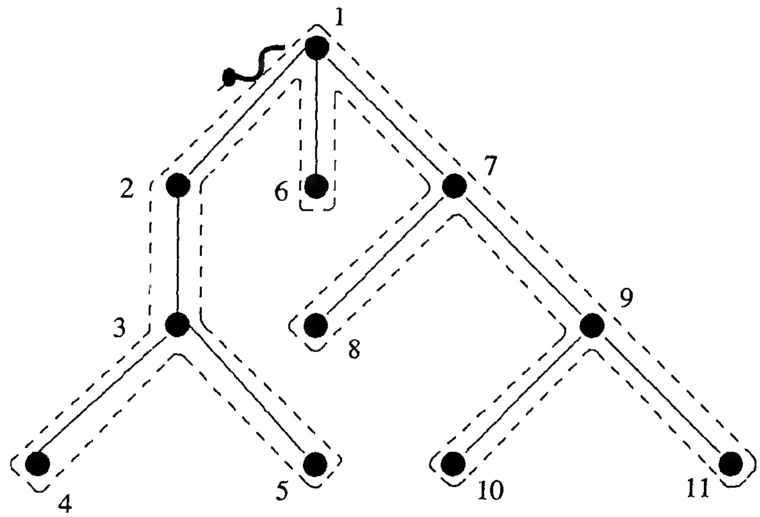


Figure 5-14. A plane tree traversed in preorder.

the preorder. Figure 5-14 shows the path of the worm on a plane tree τ, with the vertices labeled 1 to 11 in preorder.

Given a plane forest σ, let $\operatorname{ord}(\sigma) = (v_1, \ldots, v_n)$, and set $\delta_i = \deg v_i$ (the number of successors of v_i). Now define a word $w(\sigma) \in \mathcal{A}^*$ by

$$w(\sigma) = x_{\delta_1} x_{\delta_2} \cdots x_{\delta_n}.$$

For the forest σ of Figure 5-12 we have

$$w(\sigma) = x_2 x_3 x_0^5 x_1^2 x_2 x_0^2,$$

while for the tree τ of Figure 5-14,

$$w(\tau) = x_3 x_1 x_2 x_0^3 x_2 x_0 x_2 x_0^2.$$

The following fundamental lemma has a fairly straightforward proof by induction, which will be omitted here.

5.3.9 Lemma. *Let $w \in \mathcal{A}^*$. Then the map $\sigma \mapsto w(\sigma)$ is a bijection from the set of plane forests σ to $\mathcal{B}^*$.*

We now have all the ingredients necessary for our main result on plane forests.

5.3.10 Theorem. *Let $\mathbf{r} = (r_0, r_1, \ldots, r_m) \in \mathbb{N}^{m+1}$, with $\sum r_i = n$ and $\sum (1-i) r_i = k > 0$. Then the number $P(\mathbf{r})$ of plane forests (necessarily with n vertices and k components) of type $\mathbf{r}$ (i.e., r_i vertices have i successors) is given by*

$$P(\mathbf{r}) = \frac{k}{n} \binom{n}{r_0, r_1, \ldots, r_m}.$$

First Proof. The proof is an immediate consequence of Lemma 4.7.12, but for convenience we repeat the argument here. By Lemma 5.3.9, $P(\mathbf{r})$ is equal to the number of words $w \in \mathcal{B}^*$ with r_i x_i's for all i. (Regard $r_i = 0$ for $i > m$.) Denote by $\mathcal{B}^*_{\mathbf{r}}$ the set of all $P(\mathbf{r})$ such words, and similarly let $\mathcal{A}^*_{\mathbf{r}}$ be the set of all words in $\mathcal{A}^*$ with r_i x_i's for all i. Define a map $\psi : \mathcal{B}^*_{\mathbf{r}} \times [n] \to \mathcal{A}^*_{\mathbf{r}} \times [k]$ as follows. Let $w = w_1 w_2 \cdots w_n = u_1 u_2 \cdots u_k \in \mathcal{B}^*_{\mathbf{r}}$, where $w_i \in \mathcal{A}$ and $u_i \in \mathcal{B}$. Choose $i \in [n]$ and suppose w_i is a letter of u_j. Then set

$$\psi(w, i) = (w_i w_{i+1} \cdots w_{i-1}, j).$$

By Lemma 5.3.6 ψ is injective, while by Lemma 5.3.7 (and the fact that $\phi(w) = -k$ if $w \in \mathcal{B}^*$) ψ is surjective. Hence

$$nP(\mathbf{r}) = k(\#\mathcal{A}^*_{\mathbf{r}}).$$

But clearly by the formula for $|\mathfrak{S}(M)|$ at the end of Section 1.2 we have

$$\#\mathcal{A}^*_{\mathbf{r}} = \binom{n}{r_0, r_1, \ldots, r_m}, \tag{5.51}$$

and the proof follows. □

Second Proof. Let $w \in \mathcal{A}^*_{\mathbf{r}}$ (as defined in the first proof), and let $C(w)$ be the set of all *distinct* conjugates of w. If $\#C(w) = m$ then m divides n, and every element of $C(w)$ occurs exactly n/m times among the n conjugates of w. It follows from Lemma 5.3.6 that exactly k conjugates of w belong to $\mathcal{B}^*$. Hence there are exactly $(k/n)m$ *distinct* conjugates of w belonging to $\mathcal{B}^*$, so the total number of distinct conjugates of elements of $\mathcal{A}^*_{\mathbf{r}}$ belonging to $\mathcal{B}^*$ is $(k/n)(\#\mathcal{A}^*_{\mathbf{r}})$. The proof follows from Lemma 5.3.9 and (5.51). □

The situation of the previous proof is simplest when $k = 1$. Here the n conjugates $w_i w_{i+1} \cdots w_{i-1}$ of $w = w_1 w_2 \cdots w_n \in \mathcal{A}^*_{\mathbf{r}}$ are all distinct, and exactly one of them lies in $\mathcal{B}^*$. Thus $P(\mathbf{r}) = (1/n)(\#\mathbb{A}^*_{\mathbf{r}})$. The fact that the conjugates of w are all distinct may be seen directly from the formula $\sum(i-1)r_i = -k$, since if $w = v^p$ then $p | r_i$ for all i, so $p | k$.

5.3.11 Example. How many plane trees have three endpoints, one vertex of degree one, and two of degree two? This is the case $\mathbf{r} = (3, 1, 2)$. Since $\sum(i-1)r_i = -1 \cdot 3 + 0 \cdot 1 + 1 \cdot 2 = -1$, such trees exist; and

$$P(\mathbf{r}) = \frac{1}{6}\binom{6}{3, 1, 2} = 10,$$

in agreement with Figure 5-11.

5.3.12 Example. How many plane binary trees τ have $n+1$ endpoints? ("Binary" means here that every non-endpoint vertex has two successors. Without the adjective "plane," "binary" has a different meaning as explained in the Appendix of Volume 1.) One sees easily that τ has exactly n vertices of degree two. Hence $\mathbf{r}=(n+1,0,n)$, and

$$P(\mathbf{r})=\frac{1}{2n+1}\binom{2n+1}{n}=\frac{1}{n+1}\binom{2n}{n},$$

a Catalan number. These numbers made several appearances in Volume 1 and will be discussed in more detail in the next chapter (see in particular Exercise 6.19). Note that in the context of the second proof of Theorem 5.3.10, we obtain the expression

$$\frac{1}{2n+1}\binom{2n+1}{n}$$

because there are $\binom{2n+1}{n}$ sequences of n 1's and $n+1$ -1's, and each of them have $2n+1$ distinct conjugates, of which exactly one has all its partial sums (except for the last sum) nonnegative. Alternatively, there are $\binom{2n}{n}$ sequences of n 1's and $n+1$ -1's that end with a -1. Each of them has $n+1$ distinct conjugates beginning with a 1, of which exactly one has all partial sums nonnegative except for the last partial sum. This gives directly the expression $\frac{1}{n+1}\binom{2n}{n}$ for the number of plane binary trees with $n+1$ endpoints.

5.4 The Lagrange Inversion Formula

The set $xK[[x]]$ of all formal power series $a_1x+a_2x^2+\cdots$ with zero constant term over a field K forms a monoid under the operation of functional composition. The identity element of this monoid is the power series x. Recall from Example 5.2.5 that if $f(x)=a_1x+a_2x^2+\cdots\in K[[x]]$, then we call a power series $g(x)$ a *compositional inverse* of f if $f(g(x))=g(f(x))=x$, in which case we write $g(x)=f^{\langle -1\rangle}(x)$. The following simple proposition explains when $f(x)$ has a compositional inverse.

5.4.1 Proposition. *A power series $f(x)=a_1x+a_2x^2+\cdots\in K[[x]]$ has a compositional inverse $f^{\langle -1\rangle}(x)$ if and only if $a_1\neq 0$, in which case $f^{\langle -1\rangle}(x)$ is unique. Moreover, if $g(x)=b_1x+b_2x^2+\cdots$ satisfies either $f(g(x))=x$ or $g(f(x))=x$, then $g(x)=f^{\langle -1\rangle}(x)$.*

Proof. Assume that $g(x)=b_1x+b_2x^2+\cdots$ satisfies $f(g(x))=x$. We then have

$$a_1(b_1x+b_2x^2+b_3x^3+\cdots)+a_2(b_1x+b_2x^2+\cdots)^2+a_3(b_1x+\cdots)^3=x.$$

Equating coefficients on both sides yields the infinite system of equations

$$
\begin{aligned}
a_1 b_1 &= 1 \\
a_1 b_2 + a_2 b_1^2 &= 0 \\
a_1 b_3 + 2a_2 b_1 b_2 + a_3 b_1^3 &= 0 \\
&\vdots
\end{aligned}
$$

We can solve the first equation (uniquely) for b_1 if and only if $a_1 \neq 0$. We can then solve the second equation uniquely for b_2, the third for b_3, etc. Hence $g(x)$ exists if and only if $a_1 \neq 0$, in which case it is unique. The remaining assertions are special cases of the fact that in a group every left or right inverse is a two sided inverse. For the present situation, suppose for instance that $f(g(x)) = x$ and $h(f(x)) = x$. Substitute $g(x)$ for x in the second equation to get $h(x) = g(x)$, etc. □

In some cases the equation $y = f(x)$ can be solved directly for x, yielding $x = f^{\langle -1\rangle}(y)$. For instance, one can verify in this way that

$$
\begin{aligned}
(e^x - 1)^{\langle -1\rangle} &= \log(1+x) \\
\left(\frac{a+bx}{c+dx}\right)^{\langle -1\rangle} &= \frac{-a+cx}{b-dx} \qquad \text{if } ad \neq bc.
\end{aligned}
$$

In most cases, however, a simple explicit formula for $f^{\langle -1\rangle}(x)$ will not exist. We can still ask if there is a nice formula or combinatorial interpretation of the *coefficients* of $f^{\langle -1\rangle}(x)$. For instance, from (5.44) we have

$$
(xe^{-x})^{\langle -1\rangle} = \sum_{n\geq 1} n^{n-1}\frac{x^n}{n!}. \tag{5.52}
$$

Recall that we are always assuming that $\operatorname{char} K = 0$. With this assumption, the Lagrange inversion formula will express the coefficients of $f^{\langle -1\rangle}(x)$ in terms of coefficients of certain other power series. This will allow us to derive results such as (5.52) in a routine, systematic way. Somewhat more generally, we obtain an expression for the coefficients of $[f^{\langle -1\rangle}(x)]^k$ for any $k \in \mathbb{P}$. In effect this determines $g(f^{\langle -1\rangle}(x))$ for any $g(x)$, since if $g(x) = \sum b_k x^k$ then $g(f^{\langle -1\rangle}(x)) = \sum b_k [f^{\langle -1\rangle}(x)]^k$.

We will give three proofs of the Lagrange inversion formula. The first proof is a direct algebraic argument. The second proof regards power series as *ordinary* generating functions and is based on our enumeration of plane forests (Theorem 5.3.10). Our final proof regards power series as *exponential* generating functions and is based on our enumeration of planted forests (Theorem 5.3.4). Thus we will give two combinatorial proofs of Lagrange inversion, one using (unlabeled) plane forests and the other (labeled) planted forests.

5.4.2 Theorem (The Lagrange inversion formula). *Let* $F(x) = a_1x + a_2x^2 + \cdots \in xK[[x]]$, *where* $a_1 \neq 0$ *(and char* $K = 0$*), and let* $k, n \in \mathbb{Z}$. *Then*

$$n[x^n]F^{\langle -1\rangle}(x)^k = k[x^{n-k}]\left(\frac{x}{F(x)}\right)^n = k[x^{-k}]F(x)^{-n}. \tag{5.53}$$

(The second equality is trivial.) Equivalently, suppose $G(x) \in K[[x]]$ *with* $G(0) \neq 0$, *and let* $f(x)$ *be defined by*

$$f(x) = xG(f(x)). \tag{5.54}$$

Then

$$n[x^n]f(x)^k = k[x^{n-k}]G(x)^n. \tag{5.55}$$

Note 1. If $k < 0$ then $F^{\langle -1\rangle}(x)^k$ and $f(x)^k$ are Laurent series of the form $\sum_{i\geq k} p_i x^i$. Note also that if $n < k$ then both sides of (5.53) and (5.55) are 0.

Note 2. Equations (5.53) and (5.55) are equivalent since the statement that $f(x) = F^{\langle -1\rangle}(x)$ is easily seen to mean the same as $f(x) = xG(f(x))$ where $G(x) = x/F(x)$.

First Proof of Theorem 5.4.2. The first proof is based on the following innocuous observation: If $y = \sum_{n\in\mathbb{Z}} c_n x^n$ is a Laurent series, then

$$[x^{-1}]y' = 0, \tag{5.56}$$

i.e., the derivative of a Laurent series has no x^{-1} term.

Now set

$$F^{\langle -1\rangle}(x)^k = \sum_{i\geq k} p_i x^i,$$

so

$$x^k = \sum_{i\geq k} p_i F(x)^i.$$

Differentiate both sides to obtain

$$kx^{k-1} = \sum_{i\geq k} ip_i F(x)^{i-1}F'(x)$$

$$\Rightarrow \quad \frac{kx^{k-1}}{F(x)^n} = \sum_{i\geq k} ip_i F(x)^{i-n-1}F'(x). \tag{5.57}$$

Here we are expanding both sides of (5.57) as elements of $K((x))$, i.e., as Laurent series with finitely many negative exponents. For instance,

$$\frac{kx^{k-1}}{F(x)^n} = \frac{kx^{k-1}}{(a_1x + a_2x^2 + \cdots)^n}$$
$$= kx^{k-n-1}(a_1 + a_2x + \cdots)^{-n}.$$

We wish to take the coefficient of x^{-1} on both sides of (5.57). Since

$$F(x)^{i-n-1}F'(x) = \frac{1}{i-n}\frac{d}{dx}F(x)^{i-n}, \qquad i \neq n,$$

it follows from (5.56) that the coefficient of x^{-1} on the right-hand side of (5.57) is

$$[x^{-1}]np_nF(x)^{-1}F'(x) = [x^{-1}]np_n\left(\frac{a_1 + 2a_2x + \cdots}{a_1x + a_2x^2 + \cdots}\right)$$
$$= [x^{-1}]np_n\left(\frac{1}{x} + \cdots\right)$$
$$= np_n.$$

Hence

$$[x^{-1}]\frac{kx^{k-1}}{F(x)^n} = np_n = n[x^n]F^{\langle -1\rangle}(x)^k,$$

which is equivalent to (5.53). □

Second Proof (only for $k \geq 1$). Let $t_0, t_1, \ldots$ be (commuting) indeterminates, and set

$$G(x) = t_0 + t_1x + \cdots.$$

If σ is a plane forest, set

$$t^\sigma = \prod_{i\geq 0} t_i^{r_i(\sigma)}, \tag{5.58}$$

where $r_i(\sigma)$ is the number of vertices of σ of degree i. Now set

$$s_n = \sum_\tau t^\tau,$$

summed over all plane trees with n vertices. For instance,

$$s_1 = t_0, \qquad s_2 = t_0t_1,$$
$$s_3 = t_0t_1^2 + t_0^2t_2.$$

Let

$$f(x) = \sum_{n\geq 1} s_n x^n. \tag{5.59}$$

If τ is a plane tree whose root has j subtrees, then τ is obtained by choosing j (nonempty) plane trees, arranging them in linear order, and adjoining a root of degree j attached to the roots of the j plane trees. Thus

$$t_j x f(x)^j = \sum_{n\geq 1}\left(\sum_{\tau} t^{\tau}\right)x^n \tag{5.60}$$

where τ runs over all plane trees with n vertices whose root is of degree j. Summing over all $j \geq 1$ yields

$$xG(f(x)) = f(x). \tag{5.61}$$

Now let $k \in \mathbb{P}$. By the definition (5.59) of $f(x)$, we have

$$f(x)^k = \sum_{n\geq 1}\left(\sum_{\sigma} t^{\sigma}\right)x^n, \tag{5.62}$$

where σ runs over all plane forests with n vertices and k components. On the other hand, from Theorem 5.3.10 we have

$$[x^n]f(x)^k = \sum_{\sigma} t^{\sigma} = \frac{k}{n}\sum_{r_0,r_1,\ldots}\binom{n}{r_0, r_1, \ldots} t_0^{r_0} t_1^{r_1}\cdots,$$

summed over all $\mathbb{N}$-sequences $r_0, r_1, \ldots$ satisfying $\sum r_i = n$ and $\sum(i-1)r_i = -k$, or equivalently $\sum r_i = n$ and $\sum i r_i = n-k$. But

$$\begin{aligned} G(x)^n &= (t_0 + t_1 x + \cdots)^n \\ &= \sum_{r_0+r_1+\cdots=n}\binom{n}{r_0, r_1, \ldots} t_0^{r_0} t_1^{r_1}\cdots x^{\Sigma i r_i}. \end{aligned}$$

Thus

$$[x^n]f(x)^k = \frac{k}{n}[x^{n-k}]G(x)^n,$$

which is equivalent to (5.55). Since $G(x)$ has "general coefficients" (i.e., independent indeterminates), the proof follows. □

Note that this proof yields an explicit combinatorial formula (5.62) for the coefficients of $F^{\langle -1\rangle}(x)^k = f(x)^k$ in terms of the coefficients of $x/F(x) = G(x)$.

Third Proof of Theorem 5.4.2 (again only for $k \geq 1$). This proof is analogous to the previous proof, but instead of plane forests we use planted forests on $[n]$. Since the vertices are labeled (by elements of $[n]$), it is necessary to use exponential rather than ordinary generating functions. Thus we set

$$G(x) = \sum_{n \geq 0} t_n \frac{x^n}{n!}.$$

If σ is a planted forest on $[n]$, then let $r_i(\sigma)$ be the number of vertices of degree i, and as in (5.58) set $t^\sigma = \prod t_i^{r_i(\sigma)}$. Now set

$$s_n = \sum_\tau t^\tau,$$

summed over all rooted trees on $[n]$, and let

$$\begin{aligned} f(x) &= \sum_{n \geq 1} s_n \frac{x^n}{n!} \\ &= t_0 x + 2t_0 t_1 \frac{x^2}{2!} + \left(6t_0 t_1^2 + 3t_0^2 t_2\right) \frac{x^3}{3!} + \cdots . \end{aligned}$$

If τ is a rooted tree on $[n]$ whose root has degree k, then τ is obtained by choosing a root $r \in [n]$ and then placing k rooted trees on the remaining vertices $[n] - \{r\}$. By Proposition 5.1.3, we have

$$f(x)^k = \sum_{n \geq 1} \left(\sum_\zeta t^\zeta \right) \frac{x^n}{n!},$$

where ζ runs over all *ordered* k-tuples of rooted trees with total vertex set $[n]$. Thus (since rooted trees are nonempty, so there are $k!$ ways to order k of them on $[n]$),

$$\frac{1}{k!} f(x)^k = \sum_{n \geq 1} \left(\sum_\sigma t^\sigma \right) \frac{x^n}{n!}, \tag{5.63}$$

where σ runs over all planted forests on $[n]$ with k components. Hence by Proposition 5.1.15 (equations (5.15) and (5.19)), we have that

$$\frac{t_k}{k!} x f(x)^k = \sum_{n \geq 1} \left(\sum_\zeta t^\zeta \right) \frac{x^n}{n!},$$

where now ζ runs over all rooted trees on $[n]$ whose root has degree k. Summing over all $k \geq 1$ yields, as in (5.61), $f(x) = xG(f(x))$.

Now let $k \in \mathbb{P}$. We have from (5.63) and Corollary 5.3.5 that

$$\left[\frac{x^n}{n!}\right] \frac{1}{k!} f(x)^k = \frac{k}{n}\binom{n}{k} \sum_{r_0, r_1, \ldots} \frac{(n-k)! t_0^{r_0} t_1^{r_1} \cdots}{0!^{r_0} 1!^{r_1} \cdots} \binom{n}{r_0, r_1, \ldots},$$

summed over all $\mathbb{N}$-sequences $r_0, r_1, \ldots$ satisfying $\sum r_i = n$ and $\sum i r_i = n - k$. Equivalently,

$$[x^n] f(x)^k = \frac{k}{n} \sum_{\substack{r_0, r_1, \ldots \\ \Sigma r_i = n \\ \Sigma i r_i = n-k}} \binom{n}{r_0, r_1, \ldots} \frac{t_0^{r_0} t_1^{r_1} \cdots}{0!^{r_0} 1!^{r_1} \cdots}.$$

But

$$\begin{aligned} G(x)^n &= \left(t_0 + t_1 \frac{x}{1!} + t_2 \frac{x^2}{2!} + \cdots\right)^n \\ &= \sum_{r_0 + r_1 + \cdots = n} \binom{n}{r_0, r_1, \ldots} \frac{t_0^{r_0} t^{r_1} \cdots}{0!^{r_0} 1!^{r_1} \cdots} x^{\Sigma i r_i}. \end{aligned}$$

Thus

$$[x^n] f(x)^k = \frac{k}{n} [x^{n-k}] G(x)^n,$$

as desired. □

5.4.3 Corollary. *Preserve the notation of Theorem 5.4.2. Then for any power series $H(x) \in K[[x]]$ (or Laurent series $H(x) \in K((x))$) we have*

$$n[x^n] H(F^{\langle -1 \rangle}(x)) = [x^{n-1}] H'(x) \left(\frac{x}{F(x)}\right)^n. \tag{5.64}$$

Equivalently,

$$n[x^n] H(f(x)) = [x^{n-1}] H'(x) G(x)^n, \tag{5.65}$$

where $f(x) = xG(f(x))$.

Proof. By linearity (for *infinite* linear combinations) it suffices to prove (5.64) or (5.65) for $H(x) = x^k$. But this is equivalent to (5.53) or (5.55). □

Let us consider some simple examples of the use of the Lagrange inversion formula. Additional applications appear in the exercises.

5.4.4 Example. We certainly should be able to deduce the formula

$$(xe^{-x})^{\langle -1\rangle} = \sum_{n\geq 1} n^{n-1}\frac{x^n}{n!}$$

(equivalent to (5.44)) directly from Theorem 5.4.2. Indeed, letting $F(x)=xe^{-x}$ and $k=1$ in (5.53) gives

$$\begin{aligned}[x^n](xe^{-x})^{\langle -1\rangle} &= \frac{1}{n}[x^{n-1}]e^{nx}\\ &= \frac{1}{n}\frac{n^{n-1}}{(n-1)!} = \frac{n^{n-1}}{n!}.\end{aligned}$$

More generally, for any $k \in \mathbb{Z}$ we get

$$\begin{aligned}[x^n]\left((xe^{-x})^{\langle -1\rangle}\right)^k &= \frac{k}{n}[x^{n-k}]e^{nx}\\ &= \frac{k}{n}\cdot\frac{n^{n-k}}{(n-k)!}.\end{aligned} \tag{5.66}$$

Thus the number of k-component planted forests on $[n]$ is equal to

$$\frac{n!}{k!}\cdot\frac{k}{n}\cdot\frac{n^{n-k}}{(n-k)!} = \binom{n-1}{k-1}n^{n-k},$$

agreeing with Proposition 5.3.2. Note also that setting $k=-1$ in (5.66) yields

$$[x^n]\frac{1}{(xe^{-x})^{\langle -1\rangle}} = -\frac{n^n}{(n+1)!},\ \ n\geq -1 \qquad (\text{with } 0^0=1).$$

Hence

$$\left(\sum_{n\geq 1} n^{n-1}\frac{x^n}{n!}\right)^{-1} = -\sum_{n\geq -1} n^n\frac{x^n}{(n+1)!}.$$

A little rearranging yields the interesting identity

$$\left(1-\sum_{n\geq 1}(n-1)^{n-1}\frac{x^n}{n!}\right)^{-1} = \sum_{n\geq 0}(n+1)^{n-1}\frac{x^n}{n!}. \tag{5.67}$$

Compare Exercise 5.42.

5.4.5 Example. Let A be a subset of $\{2, 3, \ldots\}$. Let $t_A(n)$ denote the number of ways of beginning with an n-set S, then partitioning S into k blocks where $k \in A$, then partitioning each non-singleton block into k blocks where $k \in A$, etc., until only singleton blocks remain. (In particular, we can never have a block whose cardinality is strictly between 1 and $\min A$.) Set $t_A(0)=0$, and set $y=E_{t_A}(x)$. Then, as a generalization of both (5.26) and (5.28), we have

$$\sum_{n\in A}\frac{y^n}{n!}=y-x.$$

Hence

$$y=\left(x-\sum_{k\in A}\frac{x^k}{k!}\right)^{\langle -1\rangle},$$

so by Theorem 5.4.2,

$$t_A(n)=\left[\frac{x^n}{n!}\right]y=\left[\frac{x^{n-1}}{(n-1)!}\right]\left(1-\sum_{k\in A}\frac{x^{k-1}}{k!}\right)^{-n}.$$

When A consists of a single element k, then we have

$$\begin{aligned}\left(1-\frac{x^{k-1}}{k!}\right)^{-n}&=\sum_{j\geq 0}\binom{n+j-1}{j}\frac{x^{j(k-1)}}{k!^j}\\&=\sum_{j\geq 0}\binom{n+j-1}{j}\frac{(j(k-1))!}{k!^j}\frac{x^{j(k-1)}}{(j(k-1))!}.\end{aligned}$$

Thus (writing t_k for $t_{\{k\}}$) $t_k(n)=0$ unless $n=j(k-1)+1$ for some $j\in\mathbb{N}$, and

$$\begin{aligned}t_k(j(k-1)+1)&=\binom{jk}{j}\frac{(j(k-1))!}{k!^j}\\&=\frac{(jk)!}{j!\,k!^j}.\end{aligned}$$

A combinatorial proof can be given along the lines of Example 5.2.6.

5.5 Exponential Structures

There are many possible generalizations of the compositional and exponential formulas (Theorem 5.1.4 and Corollary 5.1.6). We will consider here a

generalization involving partially ordered sets much in the spirit of binomial posets (Chapter 3.15).

5.5.1 Definition. An *exponential structure* is a sequence $\mathbf{Q} = (Q_1, Q_2, \ldots)$ of posets satisfying the following three axioms:

(E1) For each $n \in \mathbb{P}$, Q_n is finite and has a unique maximal element $\hat{1}_n$ (denoted simply by $\hat{1}$), and every maximal chain of Q_n has n elements (or length $n-1$).

(E2) If $\pi \in Q_n$, then the interval $[\pi, \hat{1}]$ is isomorphic to Π_k (the lattice of partitions of $[k]$) for some k. We then write $|\pi| = k$. Thus if $|\pi| = k$, then every saturated chain from π to $\hat{1}$ has k elements.

(E3) Suppose $\pi \in Q_n$ and ρ is a minimal element of Q_n satisfying $\rho \le \pi$. Thus by (E1) and (E2), $[\rho, \hat{1}] \cong \Pi_n$. It follows from Example 3.10.4 that $[\rho, \pi] \cong \Pi_1^{a_1} \times \Pi_2^{a_2} \times \cdots \times \Pi_n^{a_n}$ for unique $a_1, a_2, \ldots, a_n \in \mathbb{N}$ satisfying $\sum i a_i = n$ (and $\sum a_i = |\pi|$). We require that the subposet $\Lambda_\pi = \{\sigma \in Q_n : \sigma \le \pi\}$ of Q_n be isomorphic to $Q_1^{a_1} \times Q_2^{a_2} \times \cdots \times Q_n^{a_n}$. In particular, if ρ' is another minimal element of Q_n satisfying $\rho' \le \pi$, then $[\rho, \pi] \cong [\rho', \pi]$. We call $(a_1, a_2, \ldots, a_n)$ the *type* of π.

Intuitively, one should think of Q_n as forming a set of "decompositions" of some structure S_n of "size" n into "pieces" that are smaller S_i's. Then (E2) states that given a decomposition of S_n, one can take any partition of the pieces of the decomposition and join together the pieces in each block in a unique way to obtain a coarser decomposition. Moreover, (E3) states that each piece can be decomposed independently to form a finer decomposition.

If $\mathbf{Q} = (Q_1, Q_2, \ldots)$ is an exponential structure, then let $M(n)$ denote the number of minimal elements of Q_n. As will be seen below, all the basic combinatorial properties of $\mathbf{Q}$ can be deduced from the numbers $M(n)$. We call the sequence $\mathbf{M} = (M(1), M(2), \ldots)$ the *denominator sequence* of $\mathbf{Q}$. $M(n)$ turns out to play a role for exponential structures analogous to that of the factorial function of a binomial poset (see Definition 3.15.2(c)).

We now proceed to some examples of exponential structures.

5.5.2 Example. (a) The prototypical example of an exponential structure is given by $Q_n = \Pi_n$. In this case we have $M(n) = 1$.
(b) Let $V_n = V_n(q)$ be an n-dimensional vector space over the finite field $\mathbb{F}_q$. Let Q_n consist of all collections $\{W_1, W_2, \ldots, W_k\}$ of subspaces of V_n such that $\dim W_i > 0$ for all i, and such that $V_n = W_1 \oplus W_2 \oplus \cdots \oplus W_k$ (direct sum). An element of Q_n is called a *direct sum decomposition* of V_n. We order Q_n in an obvious way by refinement, viz., $\{W_1, W_2, \ldots, W_k\} \le \{W_1', W_2', \ldots, W_j'\}$ if each W_r is contained in some W_s'. It is easily seen that $(Q_1, Q_2, \ldots)$ is an exponential

structure with

$$M(n) = q^{\binom{n}{2}}(\mathbf{n})!/n!,$$

where $(\mathbf{n})! = (1+q)(1+q+q^2)\cdots(1+q+\cdots+q^{n-1})$ as in Section 1.3.
(c) Let $\mathbf{Q} = (Q_1, Q_2, \ldots)$ be an exponential structure with denominator sequence $\mathbf{M} = (M(1), M(2), \ldots)$. Fix $r \in \mathbb{P}$, and define $Q_n^{(r)}$ to be the subposet of Q_{rn} consisting of all $\pi \in Q_{rn}$ of type $(a_1, a_2, \ldots, a_{rn})$ such that $a_i = 0$ unless r divides i. Then $\mathbf{Q}^{(r)} = (Q_1^{(r)}, Q_2^{(r)}, \ldots)$ is an exponential structure with denominator sequence $\mathbf{M}^{(r)} = (M_r(1), M_r(2), \ldots)$ given by

$$M_r(n) = \frac{M(rn)(rn)!}{M(r)^n\, n!\, r!^n}. \tag{5.68}$$

(Equation (5.68) can be seen by a direct argument and is also a special case of Lemma 5.5.3.) For instance, if $\mathbf{Q} = \mathbf{\Pi} = (\Pi_1, \Pi_2, \ldots)$, then $\Pi_n^{(r)}$ consists of all partitions of $[rn]$ whose block sizes are divisible by r.
(d) Let $r \in \mathbb{P}$, and let S be an n-set. An *r-partition of S* is a set

$$\pi = \{(B_{11}, B_{12}, \ldots, B_{1r}), (B_{21}, B_{22}, \ldots, B_{2r}), \ldots, (B_{k1}, B_{k2}, \ldots, B_{kr})\}$$

satisfying:

(i) For each $j \in [r]$, the set $\pi_j = \{B_{1j}, B_{2j}, \ldots, B_{kj}\}$ forms a partition of S (into k blocks), and
(ii) For fixed i, $\#B_{i1} = \#B_{i2} = \cdots = \#B_{ir}$.

The set $Q_n = Q_n(S)$ of all r-partitions of S has an obvious partial ordering by refinement which makes $(Q_1, Q_2, \ldots)$ into an exponential structure with $M(n) = n!^{r-1}$. (A minimal element ρ of $Q_n(S)$ may be identified with an $(r-1)$-tuple $(w_1, \ldots, w_{r-1})$ of permutations $w_i \in \mathfrak{S}(S)$ (the group of all permutations of the set S) via

$$\rho = \{(x, w_1(x), \ldots, w_{r-1}(x)), (y, w_1(y), \ldots, w_{r-1}(y)), \ldots\},$$

where $S = \{x, y, \ldots\}$, and where we abbreviate a one-element set $\{z\}$ as z.) The type $(a_1, a_2, \ldots, a_n)$ of $\pi \in Q_n$ is equal to the type of any of the partitions π_j, i.e., π_j has a_i blocks of size i. (By (ii), all the π_j's have the same type.)

The basic combinatorial properties of exponential structures will be obtained from the following lemma.

5.5.3 Lemma. *Let $\mathbf{Q} = (Q_1, Q_2, \ldots)$ be an exponential structure with denominator sequence $(M(1), M(2), \ldots)$. Then the number of $\pi \in Q_n$ of type $(a_1, a_2, \ldots, a_n)$ is equal to*

$$\frac{n!\, M(n)}{1!^{a_1}\cdots n!^{a_n} a_1! \cdots a_n!\, M(1)^{a_1}\cdots M(n)^{a_n}}.$$

Proof. Let $N=N(a_1,\ldots,a_n)$ be the number of pairs (ρ,π) where ρ is a minimal element of Q_n such that $\rho\le\pi$ and type $\pi=(a_1,\ldots,a_n)$. On the one hand we can pick ρ in $M(n)$ ways, and then pick $\pi\ge\rho$. The number of choices for π is the number of elements of Π_n of type $(a_1,\ldots,a_n)$, which is easily seen (e.g., by a simple variation of Proposition 1.3.2) to equal $n!/(1!^{a_1}\cdots n!^{a_n}a_1!\cdots a_n!)$. Hence

$$N=\frac{n!\,M(n)}{1!^{a_1}\cdots n!^{a_n}a_1!\cdots a_n!}. \tag{5.69}$$

On the other hand, if K is the desired number of $\pi\in Q_n$ of type $(a_1,\ldots,a_n)$, then we can pick π in K ways and then choose $\rho\le\pi$. Since Q_n has $M(n)$ minimal elements, the poset $\Lambda_\pi\cong Q_1^{a_1}\times\cdots\times Q_n^{a_n}$ has $M(1)^{a_1}\cdots M(n)^{a_n}$ minimal elements. Hence there are $M(1)^{a_1}\cdots M(n)^{a_n}$ choices for ρ, so

$$N=K\cdot M(1)^{a_1}\cdots M(n)^{a_n}. \tag{5.70}$$

The proof follows from (5.69) and (5.70). □

We come to the main result of this section.

5.5.4 Theorem (The compositional formula for exponential structures). *Let $(Q_1, Q_2,\ldots)$ be an exponential structure with denominator sequence $(M(1),M(2),\ldots)$. Given functions $f:\mathbb{P}\to K$ and $g:\mathbb{N}\to K$ with $g(0)=1$, define a new function $h:\mathbb{N}\to K$ by*

$$h(n)=\sum_{\pi\in Q_n} f(1)^{a_1}f(2)^{a_2}\cdots f(n)^{a_n}g(|\pi|),\qquad n\ge1,$$
$$h(0)=1,$$

where type $\pi=(a_1,a_2,\ldots,a_n)$ (so $|\pi|=a_1+a_2+\cdots+a_n$). Define formal power series $F,G,H\in K[[x]]$ by

$$F(x)=\sum_{n\ge1} f(n)\frac{x^n}{n!\,M(n)}$$
$$G(x)=E_g(x)=\sum_{n\ge0} g(n)\frac{x^n}{n!}$$
$$H(x)=\sum_{n\ge0} h(n)\frac{x^n}{n!\,M(n)}.$$

Then $H(x)=G(F(x))$.

Proof. By Theorem 5.1.4, we have

$$\left[\frac{x^n}{n!M(n)}\right] G(F(x)) = M(n) \sum_{\pi \in \Pi_n} \left(\frac{f(1)}{M(1)}\right)^{a_1} \cdots \left(\frac{f(n)}{M(n)}\right)^{a_n} g(|\pi|), \quad (5.71)$$

where type $\pi = (a_1, \ldots, a_n)$. Write $t(Q_n; a_1, a_2, \ldots)$ for the number of $\pi \in Q_n$ of type $(a_1, a_2, \ldots)$. By Lemma 5.5.3 we have

$$\frac{t(Q_n; a_1, a_2, \ldots)}{t(\Pi_n; a_1, a_2, \ldots)} = \frac{M(n)}{M(1)^{a_1} \cdots M(n)^{a_n}}.$$

Hence (5.71) may be rewritten

$$\left[\frac{x^n}{n!M(n)}\right] G(F(x)) = \sum_{\pi \in Q_n} f(1)^{a_1} \cdots f(n)^{a_n} g(|\pi|),$$

as desired. □

Putting $g(n) = 1$ for all $n \geq 0$ yields:

5.5.5 Corollary (The exponential formula for exponential structures). *Let $(Q_1, Q_2, \ldots)$ be an exponential structure with denominator sequence $(M(1), M(2), \ldots)$. Given a function $f : \mathbb{P} \to K$, define a new function $h : \mathbb{N} \to K$ by*

$$\begin{aligned} h(n) &= \sum_{\pi \in Q_n} f(1)^{a_1} \cdots f(n)^{a_n}, \ n \geq 1, \\ h(0) &= 1, \end{aligned}$$

where type $\pi = (a_1, \ldots, a_n)$. Define $F(x)$ and $H(x)$ as in Theorem 5.5.4. Then

$$H(x) = \exp F(x).$$

Let us turn to some examples of the use of Corollary 5.5.5.

5.5.6 Example. Let $(Q_1, Q_2, \ldots)$ be an exponential structure with denominator sequence $(M(1), M(2), \ldots)$, and write $q(n) = \#Q_n$. Letting $f(i) = 1$ for all i in Corollary 5.5.5 yields $h(n) = q(n)$, so

$$\sum_{n \geq 0} q(n) \frac{x^n}{n!M(n)} = \exp \sum_{n \geq 1} \frac{x^n}{n!M(n)}.$$

For instance, if $n!M(n) = q^{\binom{n}{2}}(\mathbf{n})!$, then by Example 5.5.2(b) we have that $q(n)$ is the number of ways to express $V_n(q)$ as a direct sum (without regard to order) of nontrivial subspaces.

More generally, let $S_{\mathbf{Q}}(n, k)$ denote the number of $\pi \in Q_n$ satisfying $|\pi| = k$ (so for $\mathbf{Q} = \mathbf{\Pi}$, $S_{\mathbf{Q}}(n, k)$ becomes the Stirling number $S(n, k)$ of the second kind).

Define a polynomial

$$W_n(t) = \sum_{\pi \in Q_n} t^{|\pi|} = \sum_{k=1}^{n} S_Q(n,k) t^k, \tag{5.72}$$

with $W_0(t) = 1$. Putting $f(i) = 1$ and $g(k) = t^k$ in Theorem 5.5.4 (or $f(i) = t$ in Corollary 5.5.5) leads to

$$\sum_{n \geq 0} W_n(t) \frac{x^n}{n!M(n)} = \exp\left(t \sum_{n \geq 1} \frac{x^n}{n!M(n)} \right), \tag{5.73}$$

which is analogous to Example 5.2.2.

5.5.7 Example. We now consider a generalization of the previous example. Let $r \in \mathbb{P}$, and define a polynomial

$$\begin{aligned} P_n(r,t) &= \sum_{\pi_1 \leq \cdots \leq \pi_r} t^{|\pi_r|}, \qquad n \geq 1, \\ P_0(r,t) &= 1, \end{aligned}$$

where the sum ranges over all r-element multichains in Q_n. In particular, $P_n(1,t) = W_n(t)$ and $P_n(r,1) = Z(Q_n, r+1)$, where $Z(Q_n, \cdot)$ is the zeta polynomial of Q_n (see Section 3.11). Now let $\bar{Q}_n$ denote Q_n with a $\hat{0}$ adjoined, and let ζ denote the zeta function of $\bar{Q}_n$ (as defined in Section 3.6). Then clearly for $n \geq 1$,

$$P_n(r,t) = \sum_{\pi \in Q_n} \left[\zeta^r(\hat{0},\pi) - \zeta^{r-1}(\hat{0},\pi) \right] t^{|\pi|}. \tag{5.74}$$

The right-hand side makes sense for any $r \in \mathbb{Z}$ and thus yields an interpretation of $P_n(r,t)$ for $r \leq 0$. In particular, since $\zeta^0(\hat{0},\pi) = 0$ for all $\pi \in Q_n$, putting $r = 0$ in (5.74) yields

$$\begin{aligned} P_n(0,t) &= -\sum_{\pi \in Q_n} \mu_n(\hat{0},\pi) t^{|\pi|} \\ &= t^{n+1} - t\chi(\bar{Q}_n, t), \end{aligned} \tag{5.75}$$

where μ_n denotes the Möbius function and χ the characteristic polynomial (as defined in Section 3.10) of $\bar{Q}_n$. Note that

$$\mu_n := \mu_n(\hat{0}, \hat{1}) = -[t] P_n(0,t) = -\frac{d}{dt} P_n(0,t)\Big|_{t=0}. \tag{5.76}$$

Now put $f(i) = P_i(r-1, 1)$ and $g(k) = t^k$ in Theorem 5.5.4 to deduce

$$\sum_{n\geq 0} P_n(r,t)\frac{x^n}{n!M(n)} = \exp\left(t\sum_{n\geq 1} P_n(r-1,1)\frac{x^n}{n!M(n)}\right)$$
$$= \left(\sum_{n\geq 0} P_n(r,1)\frac{x^n}{n!M(n)}\right)^t. \tag{5.77}$$

Note that from (5.75) we have

$$P_n(0,1) = -\sum_{\pi\in Q_n}\mu_n(\hat{0},\pi)$$
$$= \mu_n(\hat{0},\hat{0}) = 1,$$

by the recurrence (3.14) for Möbius functions. (This also follows from putting $r = 1$ in (5.77) and comparing with (5.73).) Hence setting $r = 0$ in (5.77) yields

$$\sum_{n\geq 0} P_n(0,t)\frac{x^n}{n!M(n)} = \left(\sum_{n\geq 0}\frac{x^n}{n!M(n)}\right)^t.$$

Applying d/dt to both sides and setting $t = 0$ yields from (5.77) that

$$-\sum_{n\geq 1}\mu_n\frac{x^n}{n!M(n)} = \log\sum_{n\geq 0}\frac{x^n}{n!M(n)}. \tag{5.78}$$

For instance, suppose $Q_n = \Pi_n^{(2)}$, the poset of partitions of $[2n]$ with even block sizes (Example 5.5.2(c)). By (5.68) we have $M_2(n) = (2n)!/2^n n!$. Hence

$$-\sum_{n\geq 1}\mu_n\frac{2^n x^n}{(2n)!} = \log\sum_{n\geq 0}\frac{2^n x^n}{(2n)!}.$$

Put $2x = y^2$ to obtain

$$-\sum_{n\geq 1}\mu_n\frac{y^{2n}}{(2n)!} = \log\cosh y,$$

or equivalently (by applying d/dy),

$$-\sum_{n\geq 1}\mu_n\frac{y^{2n-1}}{(2n-1)!} = \tanh y$$
$$= \sum_{n\geq 1}(-1)^{n-1}E_{2n-1}\frac{y^{2n-1}}{(2n-1)!},$$

where E_{2n-1} denotes an Euler (or tangent) number (see the end of Section 3.16).

Thus for $Q_n = \Pi_n^{(2)}$, we have

$$\mu_n = (-1)^n E_{2n-1}.$$

A primary reason for our discussion of exponential structures is to provide a general framework for extending our results on symmetric matrices with equal row and column sums (Examples 5.2.7–5.2.8) to arbitrary square matrices. (For rectangular matrices, see Exercise 5.65.) Thus let $\mathcal{M}(n, r)$ denote the set of all $n \times n$ $\mathbb{N}$-matrices $A = (A_{ij})$ for which every row and column sums to r. For instance, $\mathcal{M}(n, 0)$ consists of the $n \times n$ zero matrix, while $\mathcal{M}(n, 1)$ consists of the $n!$ $n \times n$ permutation matrices. We assume that the rows and columns of A are indexed by $[n]$. By a *k-component* of $A \in \mathcal{M}(n, r)$, we mean a pair (S, T) of nonempty subsets of $[n]$ satisfying the following two properties:

(i) $\#S = \#T = k$,
(ii) Let $A(S, T)$ be the $k \times k$ submatrix of A whose rows are indexed by S and whose columns are indexed by T, i.e., $A(S, T) = (A_{ij})$, where $(i, j) \in S \times T$. Then every row and column of $A(S, T)$ sums to r, i.e., $A(S, T) \in \mathcal{M}(k, r)$.

We call (S, T) a *component* of A if it is a k-component for some k. A component (S, T) is *irreducible* if any component (S', T') with $S' \subseteq S$ and $T' \subseteq T$ satisfies $(S', T') = (S, T)$. The matrix $A(S, T)$ is then also called *irreducible*. For instance, $(\{i\}, \{j\})$ is a 1-component (in which case it is irreducible) if and only if $A_{ij} = r$. It is easily seen that the set of irreducible components of A forms a 2-partition $\pi = \pi_A$ of $[n]$, as defined in Example 5.5.2(d). Conversely, we obtain (uniquely) a matrix $A \in \mathcal{M}(n, r)$ by choosing a 2-partition π of $[n]$ and then "attaching" an irreducible matrix to each block $(S, T) \in \pi$. There follows from Corollary 5.5.5 in the case $Q_i = \Pi_i^{(2)}$ the following result.

5.5.8 Proposition. *Let $h_r(a_1, \ldots, a_n)$ denote the number of matrices $A \in \mathcal{M}(n, r)$ such that A has a_i irreducible i-components (or equivalently, type $\pi_A = (a_1, \ldots, a_n)$). Let $f_r(n)$ be the number of irreducible $n \times n$ matrices $A \in \mathcal{M}(n, r)$. Then*

$$\sum_{n\geq 0} \sum_{a_1,\ldots,a_n} h_r(a_1, \ldots, a_n) t_1^{a_1} \cdots t_n^{a_n} \frac{x^n}{n!^2} = \exp \sum_{n\geq 1} f_r(n) t_n \frac{x^n}{n!^2}.$$

5.5.9 Corollary. (a) *Let $H(n, r) = \#\mathcal{M}(n, r)$. Then*

$$\sum_{n\geq 0} H(n, r) \frac{x^n}{n!^2} = \exp \sum_{n\geq 1} f_r(n) \frac{x^n}{n!^2}.$$

(b) *Let $H^*(n, r)$ denote the number of matrices in $\mathcal{M}(n, r)$ with no entry equal*

to r. *Then*

$$\sum_{n\geq 0} H^*(n,r)\frac{x^n}{n!^2} = \exp \sum_{n\geq 2} f_r(n)\frac{x^n}{n!^2}$$

$$= e^{-x}\sum_{n\geq 0} H(n,r)\frac{x^n}{n!^2}. \qquad (5.79)$$

Proof. (a) Put each $t_i = 1$ in Proposition 5.5.8.
(b) Since $(\{i\}, \{j\})$ is an (irreducible) 1-component if and only if $A_{ij} = r$, the proof follows by setting $t_1 = 0, t_2 = t_3 = \cdots = 1$ in Proposition 5.5.8 (and noting that $f_r(1) = 1$). □

There is a simple graph-theoretic interpretation of the 2-partition π_A associated with the matrix $A \in \mathcal{M}(n,r)$. Let $X = \{x_1, \ldots, x_n\}$ and $Y = \{y_1, \ldots, y_n\}$, and define a bipartite graph $\Gamma(A)$ with vertex bipartition (X, Y) by placing A_{ij} edges (or a single edge weighted by A_{ij}) between x_i and y_j. Thus $\Gamma(A)$ is regular of degree r. Then the connected components of $\Gamma(A)$ correspond to the irreducible components of A. More precisely, if Γ' is a connected component of Γ, then define

$$S = \{j : x_j \text{ is a vertex of } \Gamma'\}$$
$$T = \{j : y_j \text{ is a vertex of } \Gamma'\}.$$

Then (S, T) is an irreducible component of A (or block of π_A), and conversely all irreducible components of A are obtained in this way. Thus an irreducible $A \in \mathcal{M}(n,r)$ corresponds to a *connected* regular bipartite graph of degree r with $2n$ vertices. As an example, suppose

$$A = \begin{bmatrix} 0 & 2 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 \end{bmatrix}$$

The bipartite graph $\Gamma(A)$ is shown in Figure 5-15. The 2-partition π_A is given by

$$\pi = \{(234, 146), (16, 25), (5, 3)\},$$

of type $(1, 1, 1)$.

It is not difficult to compute $f_2(n)$. Indeed, an irreducible matrix $A \in \mathcal{M}(n, 2)$ is of the form $P + PQ$, where P is a permutation matrix and Q a cyclic permutation

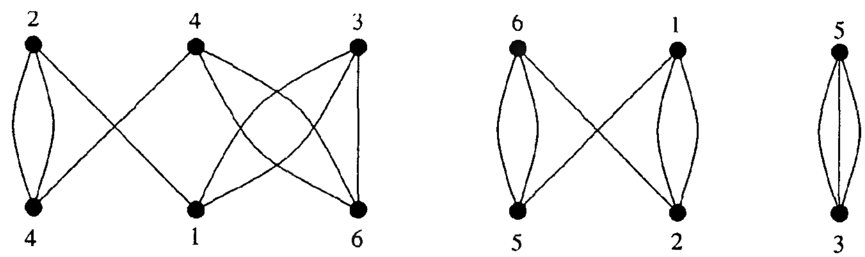


Figure 5-15. A bipartite graph $\Gamma(A)$.

matrix. In graph-theoretic terms, $\Gamma(A)$ is a connected bipartite graph of degree two (and therefore a cycle of even length ≥ 2) with vertex bipartition (X, Y) where $\#X = \#Y = n$. There are easily seen to be $\frac{1}{2}(n-1)!\,n!$ such cycles for $n \geq 2$, and of course just one for $n = 1$. (Equivalently, there are $n!$ choices for P and $(n-1)!$ choices for Q. If $n > 1$ then P and PQ could have been chosen in reverse order.) There follows from Proposition 5.5.8:

5.5.10 Proposition. *We have*

$$\sum_{n\geq 0} \sum_{a_1,\ldots,a_n} h_2(a_1,\ldots,a_n) t_1^{a_1} \cdots t_n^{a_n} \frac{x^n}{n!^2} = \exp\left(t_1 x + \frac{1}{2} \sum_{n\geq 2} t_n \frac{x^n}{n} \right). \tag{5.80}$$

5.5.11 Corollary. *We have*

$$\sum_{n\geq 0} H(n,2) \frac{x^n}{n!^2} = (1-x)^{-\frac{1}{2}} e^{\frac{1}{2}x}$$

$$\sum_{n\geq 0} H^*(n,2) \frac{x^n}{n!^2} = (1-x)^{-\frac{1}{2}} e^{-\frac{1}{2}x}. \tag{5.81}$$

Proof. Put $t_i = 1$ in (5.80) to obtain

$$\begin{aligned}
\exp\left(x + \frac{1}{2} \sum_{n\geq 2} \frac{x^n}{n} \right) &= \exp\left(\frac{1}{2} x + \frac{1}{2} \sum_{n\geq 1} \frac{x^n}{n} \right) \\
&= \exp\left(\frac{1}{2} x + \frac{1}{2} \log(1-x)^{-1} \right) \\
&= (1-x)^{-\frac{1}{2}} e^{\frac{1}{2}x}.
\end{aligned}$$

Similarly put $t_1 = 0$ and $t_2 = t_3 = \cdots = 1$ (or use (5.79) directly) to obtain (5.81). $\square$

5.6 Oriented Trees and the Matrix–Tree Theorem

A famous problem that goes back to Euler asks for what graphs G there is a closed walk that uses every edge exactly once. (There is also a version for non-closed walks.) Such a walk is called an *Eulerian tour* (also known as an *Eulerian cycle*). A graph which has an Eulerian tour is called an *Eulerian graph*. Euler's famous theorem (the first real theorem of graph theory) states that G is Eulerian if and only if it is connected (except for isolated vertices) and every vertex has even degree. Here we will be concerned with the analogous theorem for directed graphs D. We want to know not just whether an Eulerian tour exists, but how many there are. We reduce this problem to that of counting certain subtrees of D called *oriented trees*. We will prove an elegant determinantal formula for this number, and from it derive a determinantal formula, known as the *Matrix–Tree Theorem*, for the number of spanning trees of any (undirected) graph. An application of the enumeration of Eulerian tours is given to the enumeration of de Bruijn sequences. For the case of undirected graphs no analogous formula is known for the number of Eulerian tours, explaining why we consider only the directed case.

We will use the terminology and notation associated with directed graphs introduced at the beginning of Section 4.7. Let $D = (V, E, \varphi)$ be a digraph with vertex set $V = \{v_1, \dots, v_p\}$ and edge set $E = \{e_1, \dots, e_q\}$. We say that D is *connected* if it is connected as an undirected graph. A *tour* in D is a sequence $e_1, e_2, \dots, e_r$ of *distinct* edges such that the final vertex of e_i is the initial vertex of e_{i+1} for all $1 \le i \le r - 1$, and the final vertex of e_r is the initial vertex of e_1. A tour is *Eulerian* if every edge of D occurs at least once (and hence exactly once). A digraph that has no isolated vertices and contains an Eulerian tour is called an *Eulerian digraph*. Clearly an Eulerian digraph is connected. (Even more strongly, there is a directed path between any pair of vertices.) The *outdegree* of a vertex v, denoted outdeg(v), is the number of edges of G with initial vertex v. Similarly the *indegree* of v, denoted indeg(v), is the number of edges of D with final vertex v. A loop (edge of the form (v, v)) contributes one to both the indegree and outdegree. A digraph is *balanced* if indeg(v) = outdeg(v) for all vertices v.

5.6.1 Theorem. *A digraph D without isolated vertices is Eulerian if and only if it is connected and balanced.*

Proof. Assume D is Eulerian, and let $e_1, \dots, e_q$ be an Eulerian tour. As we move along the tour, whenever we enter a vertex v we must exit it, except that at the very end we enter the final vertex v of e_q without exiting it. However, at the beginning we exited v without having entered it. Hence every vertex is entered as often as it is exited and so must have the same outdegree as indegree. Therefore D is balanced, and as noted above D is clearly connected.

Now assume that D is balanced and connected. We may assume that D has at least one edge. We first claim that for any edge e of D, D has a tour (not necessarily Eulerian) for which $e = e_1$. If e_1 is a loop we are done. Otherwise we have entered

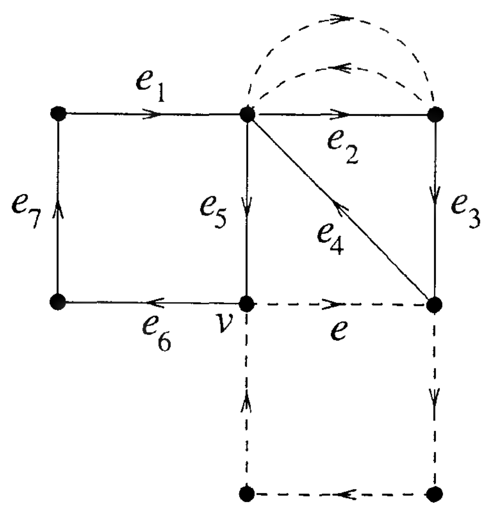


Figure 5-16. A nonmaximal tour in a balanced digraph.

the vertex $\text{fin}(e_1)$ for the first time, so since D is balanced there is some exit edge e_2. Either $\text{fin}(e_2) = \text{init}(e_1)$ and we are done, or else we have entered the vertex $\text{fin}(e_2)$ once more than we have exited it. Since D is balanced there is a new edge e_3 with $\text{fin}(e_2) = \text{init}(e_3)$. Continuing in this way, either we complete a tour or else we have entered the current vertex once more than we have exited it, in which case we can exit along a new edge. Since D has finitely many edges, eventually we must complete a tour. Thus D does have a tour for which $e = e_1$.

Now let $e_1, \ldots, e_r$ be a tour C of maximum length. We must show that $r = q$, the number of edges of D. Assume to the contrary that $r < q$. Since in moving along C every vertex is entered as often as it is exited (with $\text{init}(e_1)$ exited at the beginning and entered at the end), when we remove the edges of C from D we obtain a digraph H that is still balanced, though it need not be connected. However, since D is connected, at least one connected component H_1 of H contains at least one edge and has a vertex v in common with C. Since H_1 is balanced, there is an edge e of H_1 with initial vertex v. See Figure 5-16, where the edges of a tour C are drawn as solid lines, and the remaining edges as dotted lines. The argument of the previous paragraph shows that H_1 has a tour C' of positive length beginning with the edge e. But then when moving along C, when we reach v we can take the "detour" C' before continuing with C. This gives a tour of length longer than r, a contradiction. Hence $r = q$, and the theorem is proved. $\square$

Our primary goal is to count the number of Eulerian tours of a connected balanced digraph. A key concept in doing so is that of an oriented tree. An *oriented tree* with root v is a (finite) digraph T with v as one of its vertices, such that there is a unique directed path from any vertex u to v. In other words, for every vertex u there is a unique sequence of edges $e_1, \ldots, e_r$ such that (a) $\text{init}(e_1) = u$,

(b) $\mathrm{fin}(e_r) = v$, and (c) $\mathrm{fin}(e_i) = \mathrm{init}(e_{i+1})$ for $1 \leq i \leq r-1$. It is easy to see that this means that the underlying undirected graph (i.e., "erase" all the arrows from the edges of T) is a tree, and that all arrows in T "point toward" v. There is a surprising connection between Eulerian tours and oriented trees, given by the next result.

5.6.2 Theorem. *Let D be a connected balanced digraph with vertex set V. Fix an edge e of D, and let $v = \mathrm{init}(e)$. Let $\tau(D, v)$ denote the number of oriented (spanning) subtrees of D with root v, and let $\epsilon(D, e)$ denote the number of Eulerian tours of D starting with the edge e. Then*

$$\epsilon(D, e) = \tau(D, v) \prod_{u \in V} (\mathrm{outdeg}(u) - 1)!. \tag{5.82}$$

Proof. Let $e = e_1, e_2, \ldots, e_q$ be an Eulerian tour E in D. For each vertex $u \neq v$, let $e(u)$ be the *last exit* from u in the tour, i.e., let $e(u) = e_j$ where $\mathrm{init}(e_j) = u$ and $\mathrm{init}(e_k) \neq u$ for any $k > j$.

Claim 1. The vertices of D, together with the edges $e(u)$ for all vertices $u \neq v$, form an oriented subtree of D with root v.

Proof of Claim 1. This is a straightforward verification. Let T be the spanning subgraph of D with edges $e(u)$, $u \neq v$. Thus if $\#V = p$, then T has p vertices and $p-1$ edges. We now make the following three observations:

(a) T does not have two edges f and f' satisfying $\mathrm{init}(f) = \mathrm{init}(f')$. This is clear, since both f and f' can't be last exits from the same vertex.
(b) T does not have an edge f with $\mathrm{init}(f) = v$. This is clear, since by definition the edges of T consist only of last exits from vertices other than v, so no edge of T can exit from v.
(c) T does not have a (directed) cycle C. For suppose C were such a cycle. Let f be that edge of C which occurs after all the other edges of C in the Eulerian tour E. Let f' be the edge of C satisfying $\mathrm{fin}(f) = \mathrm{init}(f')$ $(= u$, say). We can't have $u = v$ by (b). Thus when we enter u via f, we must exit u. We can't exit u via f', since f occurs after f', in E. (Note that we cannot have $f = f'$, since then f would be a loop and therefore not a last exit.) Hence f' is not the last exit from u, contradicting the definition of T.

It is easy to see that conditions (a)–(c) imply that T is an oriented tree with root v, proving the claim.

Claim 2. We claim that the following converse to Claim 1 is true. Given a connected balanced digraph D and a vertex v, let T be an oriented (spanning) subtree

of D with root v. Then we can construct an Eulerian tour Δ as follows. Choose an edge e_1 with $\text{init}(e_1) = v$. Then continue to choose any edge possible to continue the tour, except we never choose an edge f of T unless we have to, i.e., unless it's the only remaining edge exiting the vertex at which we stand. Then we never get stuck until all edges are used, so we have constructed an Eulerian tour Δ. Moreover, the set of last exits of Δ from vertices $u \neq v$ of D coincides with the set of edges of the oriented tree T.

Proof of Claim 2. Since D is balanced, the only way to get stuck is to end up at v with no further exits available, but with an edge still unused. Suppose this is the case. At least one unused edge must be a last exit edge, i.e., an edge of T. Let u be a vertex of T closest to v in T such that the unique edge f of T with $\text{init}(f) = u$ is not in the tour. Let $y = \text{fin}(f)$. Suppose $y \neq v$. Since we enter y as often as we leave it, we don't use the last exit from y. Thus $y = v$. But then we can leave v, a contradiction. This proves Claim 2.

We have shown that every Eulerian tour Δ beginning with the edge e has associated with it a last-exit oriented subtree $T = T(\Delta)$ with root $v = \text{init}(e)$. Conversely, we have also shown that given an oriented subtree T with root v, we can obtain all Eulerian tours Δ beginning with e and satisfying $T = T(\Delta)$ by choosing for each vertex $u \neq v$ the order in which the edges from u, except the edge of T, appear in Δ. Thus for each vertex u we have $(\text{outdeg}(u) - 1)!$ choices, so for each T we have $\prod_u (\text{outdeg}(u) - 1)!$ choices. Since there are $\tau(D, v)$ choices for T, the proof follows. □

5.6.3 Corollary. *Let D be a connected balanced digraph, and let v be a vertex of D. Then the number $\tau(D, v)$ of oriented subtrees with root v is independent of v.*

Proof. Let e be an edge with initial vertex v. By equation (5.82), we need to show that the number $\epsilon(G, e)$ of Eulerian tours beginning with e is independent of e. But $e_1 e_2 \cdots e_q$ is an Eulerian tour if and only if $e_i e_{i+1} \cdots e_q e_1 e_2 \cdots e_{i-1}$ is also an Eulerian tour, and the proof follows. □

In order for Theorem 5.6.2 to be of use, we need a formula for $\tau(G, v)$. To this end, define the *Laplacian matrix* $\mathbf{L} = \mathbf{L}(D)$ of a directed graph D with vertex set $V = \{v_1, \ldots, v_p\}$ to be the $p \times p$ matrix

$$\mathbf{L}_{ij} = \begin{cases} -m_{ij} & \text{if } i \neq j \text{ and there are } m_{ij} \text{ edges with} \\ & \quad \text{initial vertex } v_i \text{ and final vertex } v_j \\ \text{outdeg}(v_i) - m_{ii} & \text{if } i = j. \end{cases}$$

Note that the diagonal entry $\text{outdeg}(v_i) - m_{ii}$ is just the number of *nonloop* edges of D with initial vertex v_i. Hence the Laplacian matrix $\mathbf{L}(D)$ is independent of

the loops of D. Note also that if every vertex of D has the same outdegree d, then the adjacency matrix $\mathbf{A}$ (defined in Section 4.7) and Laplacian matrix $\mathbf{L}$ of D are related by $\mathbf{L} = d\mathbf{I} - \mathbf{A}$, where $\mathbf{I}$ denotes the $p \times p$ identity matrix. In particular, if $\mathbf{A}$ has eigenvalues $\lambda_1, \ldots, \lambda_p$, then $\mathbf{L}$ has eigenvalues $d - \lambda_1, \ldots, d - \lambda_p$.

5.6.4 Theorem. *Let D be a loopless digraph with vertex set $V = \{v_1, \ldots, v_p\}$, and let $1 \leq k \leq p$. Let $\mathbf{L}$ be the Laplacian matrix of D, and define $\mathbf{L}_0$ to be $\mathbf{L}$ with the k-th row and column deleted. Then*

$$\det \mathbf{L}_0 = \tau(D, v_k). \tag{5.83}$$

Proof. Induction on q, the number of edges of D. First note that the theorem is true if D is not connected, since clearly $\tau(D, v_k) = 0$, while if D_1 is the component of D containing v_k and D_2 is the rest of D, then $\det \mathbf{L}_0(D) = \det \mathbf{L}_0(D_1) \cdot \det \mathbf{L}(D_2) = 0$. The least number of edges that D can have is $p-1$ (since D is connected). Suppose then that D has $p-1$ edges, so that as an undirected graph D is a tree. If D is not an oriented tree with root v_k, then some vertex $v_i \neq v_k$ of D has outdegree 0. Then $\mathbf{L}_0$ has a zero row, so $\det \mathbf{L}_0 = 0 = \tau(D, v_k)$. If on the other hand D is an oriented tree with root v_k, then there is an ordering of the set $V - \{v_k\}$ so that $\mathbf{L}_0$ is upper triangular with 1's on the main diagonal. Hence $\det \mathbf{L}_0 = 1 = \tau(D, v_k)$.

Now suppose that D has $q > p - 1$ edges, and assume the theorem for digraphs with at most $q - 1$ edges. We may assume that no edge f of D has initial vertex v_k, since such an edge belongs to no oriented tree with root v_k and also makes no contribution to $\mathbf{L}_0$. It then follows, since D has at least p edges, that there exists a vertex $u \neq v_k$ of D of outdegree at least two. Let e be an edge with $\mathrm{init}(e) = u$. Let D_1 be D with the edge e removed. Let D_2 be D with all edges e' removed such that $\mathrm{init}(e) = \mathrm{init}(e')$ and $e' \neq e$. (Note that D_2 is strictly smaller than D, since $\mathrm{outdeg}(u) \geq 2$.) By induction, we have $\det \mathbf{L}_0(D_1) = \tau(D_1, v_k)$ and $\det \mathbf{L}_0(D_2) = \tau(D_2, v_k)$. Clearly $\tau(D, v_k) = \tau(D_1, v_k) + \tau(D_2, v_k)$, since in an oriented tree T with root v_k there is exactly one edge whose initial vertex coincides with that of e. On the other hand, it follows immediately from the multilinearity of the determinant that

$$\det \mathbf{L}_0(D) = \det \mathbf{L}_0(D_1) + \det \mathbf{L}_0(D_2).$$

From this the proof follows by induction. □

The operation of removing a row and column from $\mathbf{L}(D)$ may seem somewhat contrived. In the case when D is balanced (so $\tau(D, v)$ is independent of v), we would prefer a description of $\tau(D, v)$ directly in terms of $\mathbf{L}(D)$. Such a description will follow from the next lemma.

5.6.5 Lemma. *Let $\mathbf{M}$ be a $p \times p$ matrix (with entries in a field) such that the sum of the entries in every row and column is 0. Let $\mathbf{M}_0$ be the matrix obtained*

from $\mathbf{M}$ *by removing the* i*-th row and* j*-th column. Then the coefficient of* x *in the characteristic polynomial* $\det(\mathbf{M} - x\mathbf{I})$ *of* $\mathbf{M}$ *is equal to* $(-1)^{i+j+1} p \cdot \det(\mathbf{M}_0)$. *(Moreover, the constant term of* $\det(\mathbf{M} - x\mathbf{I})$ *is* 0.*)*

Proof. The constant term of $\det(\mathbf{M} - x\mathbf{I})$ is $\det(\mathbf{M})$, which is 0 because the rows of $\mathbf{M}$ sum to 0.

For definiteness we prove the rest of the lemma only for removing the last row and column, though the proof works just as well for any row and column. Add all the rows of $\mathbf{M} - x\mathbf{I}$ except the last row to the last row. This doesn't affect the determinant, and will change the entries of the last row all to $-x$ (since the rows of $\mathbf{M}$ sum to 0). Factor out $-x$ from the last row, yielding a matrix $\mathbf{N}(x)$ satisfying $\det(\mathbf{M} - x\mathbf{I}) = -x \det \mathbf{N}(x)$. Hence the coefficient of x in $\det(\mathbf{M} - x\mathbf{I})$ is given by $-\det \mathbf{N}(0)$. Now add all the columns of $\mathbf{N}(0)$ except the last column to the last column. This does not effect $\det \mathbf{N}(0)$. Because the columns of $\mathbf{M}$ sum to 0, the last column of $\mathbf{N}(0)$ becomes the column vector $[0, 0, \ldots, 0, p]^t$. Expanding the determinant by the last column shows that $\det \mathbf{N}(0) = p \cdot \det \mathbf{M}_0$, and the proof follows. □

Suppose that the eigenvalues of the matrix $\mathbf{M}$ of Lemma 5.6.5 are equal to $\mu_1, \ldots, \mu_p$ with $\mu_p = 0$. Since $\det(\mathbf{M} - x\mathbf{I}) = -x \prod_{j=1}^{p-1}(\mu_j - x)$, we see that

$$(-1)^{i+j+1} p \cdot \det \mathbf{M}_0 = -\mu_1 \cdots \mu_{p-1}. \tag{5.84}$$

This equation allows Theorem 5.6.4, in the case of balanced digraphs, to be restated as follows.

5.6.6 Corollary. *Let* D *be a balanced digraph with* p *vertices and with Laplacian matrix* $\mathbf{L}$. *Suppose that the eigenvalues of* $\mathbf{L}$ *are* $\mu_1, \ldots, \mu_p$ *with* $\mu_p = 0$. *Then for any vertex* v *of* D,

$$\tau(D, v) = \frac{1}{p}\mu_1 \cdots \mu_{p-1}.$$

Combining Theorems 5.6.2 and 5.6.4 yields a formula for the number of Eulerian tours in a balanced digraph.

5.6.7 Corollary. *Let* D *be a connected balanced digraph with* p *vertices. Let* e *be an edge of* D. *Then the number* $\epsilon(D, e)$ *of Eulerian tours of* D *with first edge* e *is given by*

$$\epsilon(D, e) = (\det \mathbf{L}_0(D)) \prod_{u \in V} (\mathrm{outdeg}(u) - 1)!.$$

Equivalently (using Corollary 5.6.6), if $\mathbf{L}(D)$ *has eigenvalues* $\mu_1, \ldots, \mu_p$ *with*

$\mu_p = 0$, *then*

$$\epsilon(D, e) = \frac{1}{p}\mu_1 \cdots \mu_{p-1} \prod_{u \in V}(\operatorname{outdeg}(u) - 1)!.$$

Let us consider an important special case of Corollary 5.6.7. The *Laplacian matrix* $\mathbf{L} = \mathbf{L}(G)$ of the *undirected* graph G with vertex set $V = \{v_1, \ldots, v_p\}$ is the $p \times p$ matrix

$$\mathbf{L}_{ij} = \begin{cases} -m_{ij} & \text{if } i \neq j \text{ and there are } m_{ij} \text{ edges between} \\ & \quad \text{vertices } v_i \text{ and } v_j \\ \deg(v_i) - m_{ii} & \text{if } i = j, \end{cases}$$

where $\deg(v_i)$ denotes the degree (number of incident edges) of v_i. Let $\hat{G}$ be the digraph obtained from G by replacing each edge $e = uv$ of G with a pair of directed edges $u \to v$ and $v \to u$. Clearly $\hat{G}$ is balanced, and $\hat{G}$ is connected whenever G is. Choose a vertex v of G. There is an obvious one-to-one correspondence between spanning trees T of G and oriented spanning trees $\hat{T}$ of $\hat{G}$ with root v, namely, direct each edge of T toward v. Moreover, $\mathbf{L}(G) = \mathbf{L}(\hat{G})$. Let $c(G)$ denote the number of spanning trees (or *complexity*) of G. Then as an immediate consequence of Theorem 5.6.4 we obtain the following determinantal formula for $c(G)$. This formula is known as the *Matrix–Tree Theorem.*

5.6.8 Theorem (The Matrix–Tree Theorem). *Let G be a finite connected p-vertex graph without loops, with Laplacian matrix $\mathbf{L} = \mathbf{L}(G)$. Let $1 \leq i \leq p$, and let $\mathbf{L}_0$ denote $\mathbf{L}$ with the i-th row and column removed. Then*

$$c(G) = \det \mathbf{L}_0.$$

Equivalently, if $\mathbf{L}$ has eigenvalues $\mu_1, \ldots, \mu_p$ with $\mu_p = 0$, then

$$c(G) = \frac{1}{p}\mu_1 \cdots \mu_{p-1}.$$

Let us look at some examples of the use of the results we have just proved.

5.6.9 Example. Let $G = K_p$, the complete graph on p vertices. We have $\mathbf{L}(K_p) = p\mathbf{I} - \mathbf{J}$, where $\mathbf{J}$ is the $p \times p$ matrix of all 1's, and $\mathbf{I}$ is the $p \times p$ identity matrix. Since $\mathbf{J}$ has rank one, $p - 1$ of its eigenvalues are equal to 0. Since tr $\mathbf{J}$ $= p$, the other eigenvalue is equal to p. (Alternatively, the column vector of all 1's is an eigenvector with eigenvalue p.) Hence the eigenvalues of $p\mathbf{I} - \mathbf{J}$ are

p ($p-1$ times) and 0 (once). By the Matrix–Tree Theorem we get

$$c(K_p) = \frac{1}{p}p^{p-1} = p^{p-2},$$

agreeing with the formula for $t(n)$ in Proposition 5.3.2.

5.6.10 Example. Let Γ be the group $(\mathbb{Z}/2\mathbb{Z})^n$ of n-tuples of 0's and 1's under componentwise addition modulo 2. Define a "scalar product" $\alpha \cdot \beta$ on Γ by

$$(a_1, \ldots, a_n) \cdot (b_1, \ldots, b_n) = \sum a_i b_i \in \mathbb{Z}/2\mathbb{Z}.$$

Note that since $(-1)^m$ depends only on the value of the integer m modulo 2, such expressions as $(-1)^{\alpha\cdot\beta+\gamma\cdot\delta}$ are well defined for $\alpha, \beta, \gamma, \delta \in \Gamma$ whether we interpret the addition in the exponent as taking place in $\mathbb{Z}/2\mathbb{Z}$ or in $\mathbb{Z}$. In particular, there continues to hold the law of exponents $(-1)^{\alpha+\beta} = (-1)^{\alpha}(-1)^{\beta}$. Let C_n be the graph whose vertices are the elements of Γ, with two vertices α and β connected by an edge whenever $\alpha + \beta$ has exactly one component equal to 1. Thus C_n may be regarded as the graph formed by the vertices and edges of an n-dimensional cube. Equivalently, C_n is the Hasse diagram of the boolean algebra B_n, regarded as a graph. Let V be the vector space of all functions $f : \Gamma \to \mathbb{Q}$. Define a linear transformation $\Phi : V \to V$ by

$$(\Phi f)(\alpha) = nf(\alpha) - \sum_{\beta} f(\beta),$$

where β ranges over all elements of Γ adjacent to α in C_n. Note that the matrix of Φ with respect to some ordering of the basis Γ of V is just the Laplacian matrix $\mathbf{L}(C_n)$ (with respect to the same ordering of the vertices of C_n). Now for each $\gamma \in \Gamma$ define a function $\chi_\gamma \in V$ by

$$\chi_\gamma(\alpha) = (-1)^{\alpha\cdot\gamma}.$$

Then

$$(\Phi\chi_\gamma)(\alpha) = n(-1)^{\alpha\cdot\gamma} - \sum_{\beta}(-1)^{\beta\cdot\gamma},$$

with β as above. If γ has exactly k 1's, then for exactly $n-k$ values of β do we have $\beta\cdot\gamma = \alpha\cdot\gamma$, while for the remaining k values of β we have $\beta\cdot\gamma = \alpha\cdot\gamma + 1$. Hence

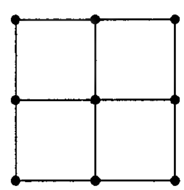
$$\begin{aligned}(\Phi\chi_\gamma)(\alpha) &= (n - [(n-k) - k])\,(-1)^{\alpha\cdot\gamma} \\ &= 2k\,\chi_\gamma(\alpha).\end{aligned}$$

It follows that χ_Y is an eigenvector of Φ with eigenvalue $2k$. It is easy to see that the χ_Y's are linearly independent, so we have found all 2^n eigenvalues of $\mathbf{L}$, viz., $2k$ is an eigenvalue of multiplicity $\binom{n}{k}$, $0 \le k \le n$. Hence from the Matrix–Tree Theorem there follows the remarkable result

$$\begin{aligned} c(C_n) &= \frac{1}{2^n}\prod_{k=1}^{n}(2k)^{\binom{n}{k}} \\ &= 2^{2^n-n-1}\prod_{k=1}^{n}k^{\binom{n}{k}}. \end{aligned} \tag{5.85}$$

A direct combinatorial proof of this formula is not known.

5.6.11 Example (The efficient mail carrier). A mail carrier has an itinerary of city blocks to which he must deliver mail. He wants to accomplish this by walking along each block twice, once in each direction, thus passing along houses on each side of the street. The blocks form the edges of a graph G, whose vertices are the intersections. The mail carrier wants simply to walk along an Eulerian tour in the digraph $\hat{G}$ defined after Corollary 5.6.7. Making the plausible assumption that the graph is connected, not only does an Eulerian tour always exist, but we can tell the mail carrier how many there are. Thus he will know how many different routes he can take to avoid boredom. For instance, suppose G is the 3×3 grid illustrated below:



This graph has 192 spanning trees. Hence the number of mail carrier routes beginning with a fixed edge (in a given direction) is $192 \cdot 1!^4 2!^4 3! = 18432$. The total number of routes is thus 18432 times twice the number of edges, viz., $18432 \times 24 = 442368$. Assuming the mail carrier delivered mail 250 days a year, it would be 1769 years before he would have to repeat a route!

5.6.12 Example (Binary de Bruijn sequences). A *binary sequence* is just a sequence of 0's and 1's. A *(binary) de Bruijn sequence* of degree n is a binary sequence $A = a_1a_2\cdots a_{2^n}$ such that every binary sequence $b_1\cdots b_n$ of length n occurs exactly once as a "circular factor" of A, i.e., as a sequence $a_ia_{i+1}\cdots a_{i+n-1}$, where the subscripts are taken modulo n if necessary. Note that there are exactly 2^n binary sequences of length n, so the only possible length of a de Bruijn sequence of degree n is 2^n. Clearly any conjugate (cyclic shift) $a_ia_{i+1}\cdots a_{2^n}a_1a_2\cdots a_{i-1}$ of a de Bruijn sequence $a_1a_2\cdots a_{2^n}$ is also a de Bruijn sequence, and we call two such

sequences *equivalent*. This relation of equivalence is obviously an equivalence relation, and every equivalence class contains exactly one sequence beginning with n 0's. Up to equivalence, there is one de Bruijn sequence of degree two, namely, 0011. It's easy to check that there are two inequivalent de Bruijn sequences of degree three, namely, 00010111 and 00011101. However, it's not clear at this point whether de Bruijn sequences exist for all n. By a clever application of Theorems 5.6.2 and 5.6.4, we will not only show that such sequences exist for all positive integers n, but will also count them. It turns out that there are *lots* of them. For instance, the number of inequivalent de Bruijn sequences of degree eight is equal to

$$1329227995784915872903807060280344576.$$

Our method of enumerating de Bruijn sequence will be to set up a correspondence between them and Eulerian tours in a certain directed graph D_n, the *de Bruijn graph* of degree n. The graph D_n has 2^{n-1} vertices, which we will take to consist of the 2^{n-1} binary sequences of length $n-1$. A pair $(a_1a_2\cdots a_{n-1}, b_1b_2\cdots b_{n-1})$ of vertices forms an edge of D_n if and only if $a_2a_3\cdots a_{n-1}=b_1b_2\cdots b_{n-2}$, i.e., e is an edge if the last $n-2$ terms of init(e) agree with the first $n-2$ terms of fin(e). Thus every vertex has indegree two and outdegree two, so D_n is balanced. The number of edges of D_n is 2^n. Moreover, it's easy to see that D_n is connected (see Lemma 5.6.13). The graphs D_3 and D_4 are shown in Figure 5-17.

Suppose that $E=e_1e_2\cdots e_{2^n}$ is an Eulerian tour in D_n. If fin(e_i) is the binary sequence $a_{i1}a_{i2}\cdots a_{i,n-1}$, then replace e_i in E by the last bit $a_{i,n-1}$. It is easy to see that the resulting sequence $\beta(E)=a_{1,n-1}a_{2,n-1}\cdots a_{2^n,n-1}$ is a de Bruijn sequence, and conversely every de Bruijn sequence arises in this way. In particular, since D_n is balanced and connected, there exists at least one de Bruijn sequence. In order to count all such sequences, we need to compute $\det \mathbf{L_0}(D_n)$. One way to do this is by a clever but messy sequence of elementary row and column operations which transforms the determinant into triangular form. We will give instead an elegant computation of the eigenvalues of $\mathbf{L}(D_n)$ (and hence of $\det \mathbf{L_0}$) based on the following simple lemma.

5.6.13 Lemma. *Let u and v be any two vertices of D_n. Then there is a unique (directed) walk from u to v of length $n-1$.*

Proof. Suppose $u=a_1a_2\cdots a_{n-1}$ and $v=b_1b_2\cdots b_{n-1}$. Then the unique path of length $n-1$ from u to v has vertices

$$a_1a_2\cdots a_{n-1},\ a_2a_3\cdots a_{n-1}b_1,\ a_3a_4\cdots a_{n-1}b_1b_2,$$
$$\ldots, a_{n-1}b_1\cdots b_{n-2},\ b_1b_2\cdots b_{n-1}. \qquad \square$$

5.6.14 Lemma. *The eigenvalues of $\mathbf{L}(D_n)$ are 0 (with multiplicity one) and 2 (with multiplicity $2^{n-1}-1$).*

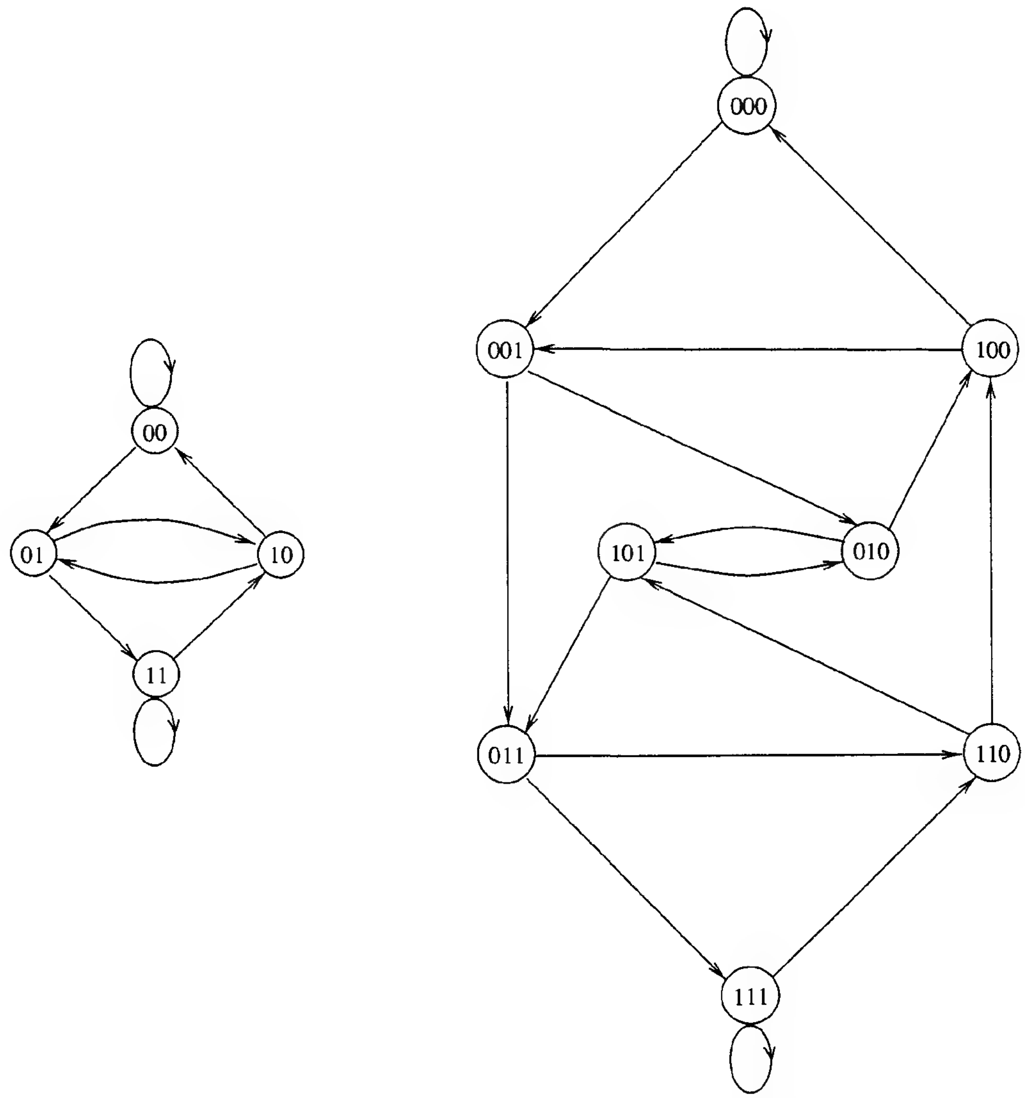


Figure 5-17. The de Bruijn graphs D_3 and D_4.

Proof. Let $\mathbf{A}(D_n)$ denote the directed adjacency matrix of D_n, i.e., the rows and columns are indexed by the vertices, with

$$\mathbf{A}_{uv} = \begin{cases} 1 & \text{if } (u, v) \text{ is an edge,} \\ 0 & \text{otherwise.} \end{cases}$$

Now Lemma 5.6.13 is equivalent to the assertion that $\mathbf{A}^{n-1} = \mathbf{J}$, the $2^{n-1} \times 2^{n-1}$ matrix of all 1's. If the eigenvalues of $\mathbf{A}$ are $\lambda_1, \ldots \lambda_{2^{n-1}}$, then the eigenvalues of $\mathbf{J} = \mathbf{A}^{n-1}$ are $\lambda_1^{n-1}, \ldots, \lambda_{2^{n-1}}^{n-1}$. By Example 5.6.9, the eigenvalues of $\mathbf{J}$ are 2^{n-1} (once) and 0 ($2^{n-1} - 1$ times). Hence the eigenvalues of $\mathbf{A}$ are 2ζ (once, where ζ is an $(n-1)$-st root of unity to be determined), and 0 ($2^{n-1} - 1$ times). Since the trace of $\mathbf{A}$ is 2, it follows that $\zeta = 1$, and we have found all the eigenvalues of A.

Now $\mathbf{L}(D_n) = 2\mathbf{I} - \mathbf{A}(D_n)$. Hence the eigenvalues of $\mathbf{L}$ are $2-\lambda_1, \ldots, 2-\lambda_{2^{n-1}}$, and the proof follows from the above determination of $\lambda_1, \ldots, \lambda_{2^{n-1}}$. □

5.6.15 Corollary. *The number $B_0(n)$ of de Bruijn sequences of degree n beginning with n 0's is equal to $2^{2^{n-1}-n}$. The total number $B(n)$ of de Bruijn sequences of degree n is equal to $2^{2^{n-1}}$.*

Proof. By the above discussion, $B_0(n)$ is the number of Eulerian tours in D_n whose first edge is the loop at vertex $00\cdots 0$. Moreover, the outdegree of every vertex of D_n is two. Hence by Corollary 5.6.7 and Theorem 5.6.14 we have

$$B_0(n) = \frac{1}{2^{n-1}} 2^{2^{n-1}-1} = 2^{2^{n-1}-n}.$$

Finally, $B(n)$ is obtained from $B_0(n)$ by multiplying by the number 2^n of edges, and the proof follows. □

Notes[1]

The compositional formula (Theorem 5.1.4) and the exponential formula (Corollary 5.1.6) had many precursors before blossoming into their present form. A purely formal formula for the coefficients of the composition of two exponential generating functions goes back to Faà di Bruno [**23**][**24**] in 1855 and 1857, and is known as *Faà di Bruno's formula*. For additional references on this formula, see [**2.3**, p. 137]. An early precursor of the exponential formula is due to Jacobi [**38**]. The idea of interpreting the coefficients of $e^{F(x)}$ combinatorially was considered in certain special cases by Touchard [**69**] and by Riddell and Uhlenbeck [**56**]. Touchard was concerned with properties of permutations and obtained our equation (5.30), from which he derived many consequences. Equation (5.30) was earlier obtained by Pólya [**50**, Sect. 13], but he was not interested in general combinatorial applications. It is also apparent from the work of Frobenius (see [**27**, bottom of p. 152 of GA]) and Hurwitz [**37**, §4] that they were aware of (5.30), even if they did not state it explicitly. Riddell and Uhlenbeck, on the other hand, were concerned with graphical enumeration and obtained our Example 5.2.1 and related results.

It was not until the early 1970s that a general combinatorial interpretation of $e^{F(x)}$ was developed independently by Foata and Schützenberger [**26**], Bender and Goldman [**3.3**], and Doubilet, Rota, and Stanley [**3.12**]. The approach most like the one taken here is that of Foata and Schützenberger. Doubilet, Rota, and Stanley use an incidence-algebra approach and prove a result (Theorem 5.1) equivalent to our Theorem 5.1.11. The most sophisticated combinatorial theory of power

[1] A reference such as [***m.n***] refers to reference n of the Notes section to Chapter m. A reference without a prefix refers to the reference list of this chapter (which follows these notes).

series composition is the theory of *species*, which is based on category theory and which was developed after the above three references by A. Joyal [**3.23**] and his collaborators. For further information on species, see [**2**]. Another category theory approach to the exponential formula was given by A. W. M. Dress and T. Müller [**16**]. The exponential formula has been frequently rediscovered in various guises; an interesting example is [**51**]. A q-analogue has been given by Gessel [**29**].

Let us turn to the applications of the exponential formula given in Section 5.2. Example 5.2.3 first appeared in [**4.36**, Example 6.6]. The generating functions (5.27) and (5.28) for total partitions and binary total partitions, as well as the explicit formula $b(n) = 1 \cdot 3 \cdot 5 \cdots (2n-3)$, are given by E. Schröder [**60**] as the fourth and third problems of his famous "*vier combinatorische Probleme*." (We will discuss the first two problems in Chapter 6.) A minor variation of the combinatorial proof given here of the formula for $b(n)$ appears in [**21**, Cor. 2], though there may be earlier proofs of a similar nature. See Exercise 5.43 for a generalization and further references. For further work related to Schröder's fourth problem, see the solution to Exercise 5.40. The generating functions and recurrence relations for $S_n(2)$ and $S_n^*(2)$ in Examples 5.2.7 and 5.2.8 were found (with a different proof from ours) by H. Gupta [**35**, (6.3), (6.4), (6.7), and (6.8)]. For a generalization, see R. Grimson [**34**]. Example 5.2.9 is due to I. Schur [**61**] and is also discussed in [**53**, Problem VII.45]. Schur considers some variants, one of which leads to the generating function for $T_n(2)$ given in Example 5.2.8. On the other hand, the generating function for $T_n^*(2)$ (equation (5.29)) essentially appears (again in a different context, discussed here in Exercise 5.23) in [**19**].

We already mentioned that equation (5.30) is due to Pólya (or possibly Frobenius or Hurwitz). It seems clear from the work of Touchard [**69**] that he was aware of the generating function $\exp \sum_{d|r} (x^d/d)$ of Example 5.2.10. The first explicit statement is due to Chowla, Herstein, and Scott [**11**], the earlier cases $r = 2$ and r prime having been investigated by Chowla, Herstein, and Moore [**10**] and by Jacobstahl [**39**], respectively. Comtet [**2.3**, Exer. 9, p. 257] discusses this subject and gives some additional references. For a significant generalization, see Exercise 5.13(a).

Example 5.2.11 was found in collaboration with I. Gessel. Similar arguments appear in Exercise 5.21 and in the paper [**48**] of Metropolis and Rota.

The concept of tree as a formal mathematical object goes back to Kirchhoff and von Staudt. Trees were first extensively investigated by Cayley, to whom the term "tree" is due. In particular, in [**9**] Cayley states the formula $t(n) = n^{n-2}$ for the number of free trees on an n-element vertex set, and he gives a vague idea of a combinatorial proof. Cayley pointed out, however, that an equivalent result had been proved earlier by Borchardt [**4**]. Moreover, this result appeared even earlier in a paper of Sylvester [**68**]. Undoubtedly Cayley and Sylvester could have furnished a complete, rigorous proof had they the inclination to do so. The first explicit combinatorial proof of the formula $t(n) = n^{n-2}$ is due to Prüfer [**52**], and is essentially the same as the case $k = 1$ of our first proof of Proposition 5.3.2. The second proof of Proposition 5.3.2 (or more precisely, the version given for

trees at the beginning of the proof) is due to Joyal [**3.23**, Example 12, pp. 15–16]. The more general formula for $p_S(n)$ given in Proposition 5.3.2 was also stated by Cayley and is implicit in the work of Borchardt. Raney [**55**] uses a straightforward generalization of Prüfer sequences to give a formal solution to the functional equation

$$\sum_i A_i e^{B_i x} = x.$$

A less obvious generalization of Prüfer sequences was given by Knuth [**40**] and is also discussed in [**47**, §2.3].

The connection between Prüfer sequences and degree sequences of trees was observed by Neville [**49**]. It was also pointed out by Moon [**45**][**46**, p. 72] and Riordan [**57**], who noted that it implied the case $k = 1$ of Theorem 5.3.4. The second proof of Theorem 5.3.4 is based on the paper [**42**] of Labelle.

The enumeration of plane (or ordered) trees by degree sequences (the case $k = 1$ of Theorem 5.3.10) is due to Erdélyi and Etherington [**20**]; their basic tool is essentially the Lagrange inversion formula. (Erdélyi and Etherington work with "non-associative combinations" rather than trees, but in [**22**] Etherington points out the connection, known to Cayley, between non-associative combinations and plane trees.) The first combinatorial proof of Theorem 5.3.10, essentially the proof given here, is due to Raney [**54**, Thm. 2.2]. (Raney works with "words" or more generally "lists of words" rather than trees; his words are essentially the Łukasiewicz words of equation (5.50).) Raney used his result to give a combinatorial proof of the Lagrange inversion formula, as discussed below. The crucial combinatorial result on which the proof of Theorem 5.3.10 is based is Lemma 5.3.7. This result (including the statement after Example 5.3.8 that if $\phi(w) = -k$ then precisely k cyclic shifts of w belong to $\mathcal{B}^*$) is part of a circle of results known as the *Cycle Lemma*. The first such result (which includes the case $\mathcal{A} = \{x_0, x_{-1}\}$ of Lemma 5.3.7) is due to Dvoretsky and Motzkin [**17**]. For further information and references, see [**18**]. For further information on the extensively developed subject of tree enumeration, see for instance [**33**][**41**, §2.3][**46**][**47**].

The Lagrange inversion formula (Theorem 5.4.2) is due, logically enough, to Lagrange [**43**]. His proof is the same as our first proof. This proof is repeated by Bromwich [**5**, Ch. VIII, §55.1], who gives many interesting applications (see our Exercises 5.53, 5.54, and 5.57). The first combinatorial proof is due to Raney [**54**]. His proof is essentially the same as our second proof, though as mentioned earlier he worked entirely with words and only implicitly with plane trees and forests. Streamlined versions of Raney's proof appear in Schützenberger [**62**] and Lothaire [**4.21**, Ch. 11]. Our third proof of Theorem 5.4.2 is essentially the same as that of Labelle [**42**]. For some further references, see [**2.3**, pp. 148–149] and [**28**].

There have been many generalizations of the Lagrange inversion formula. For fascinating surveys of multivariable Lagrange inversion formulas and their interconnections, see Gessel [**30**] and Henrici [**36**]. Gessel gives a combinatorial proof

which generalizes our third proof of Theorem 5.4.2. There has also been considerable work on q-analogues of the Lagrange inversion formula. Special cases were found by Jackson and Carlitz, followed by more general versions and/or applications due to Andrews, Cigler, Garsia, Garsia and Remmel, Gessel, Gessel and Stanton, Hofbauer, Krattenthaler, Paule, *et al.* A survey of these results is given by Stanton [**66**]. A subsequent unified approach to q-Lagrange inversion was given by Singer [**63**]. Finally, Gessel [**28**] gives a generalization of Lagrange inversion to noncommutative power series (as well as a q-analogue).

Exponential structures (Definition 5.5.1) were created by Stanley [**65**]. Their original motivation was to "explain" the formula $\mu_n = (-1)^n E_{2n-1}$ of Example 5.5.7, which had earlier been obtained by G. Sylvester [**67**] by *ad hoc* reasoning. (An equivalent result, though not stated in terms of posets and Möbius functions, had earlier been given by Rosen [**59**, Lemma 3].) Exponential structures are closely related to the exponential prefabs of Bender and Goldman [**3.3**]; see [**65**] for further information.

We have already encountered the function $H(n, r)$ of Corollaries 5.5.9 and 5.5.11 in Section 4.6 (where it was denoted $H_n(r)$). In that section we were concerned with the behavior of $H(n, r)$ for fixed n, while here we are concerned with fixed r. Corollary 5.5.11 was first proved by Anand, Dumir, and Gupta [**4.1**, §8] using a different technique (viz., first obtaining a recurrence relation). The approach we have taken here first appeared in [**4.36**, Example 6.11].

The characterization of Eulerian digraphs given by Theorem 5.6.1 is a result of Good [**32**], while the fundamental connection between oriented subtrees and Eulerian tours in a balanced digraph that was used to prove Theorem 5.6.2 was shown by van Aardenne-Ehrenfest and de Bruijn [**1**, Thm. 5a]. This result is sometimes called the BEST Theorem, after de **B**ruijn, van Aardenne-**E**hrenfest, **S**mith, and **T**utte. However, Smith and Tutte were not involved in the original discovery. (In [**64**] Smith and Tutte give a determinantal formula for the number of Eulerian tours in a special class of balanced digraphs. Van Aardenne-Ehrenfest and de Bruijn refer to the paper of Smith and Tutte in a footnote added in proof.) The determinantal formula for the number of oriented subtrees of a directed graph (Theorem 5.6.4) is due to Tutte [**70**, Thm. 3.6]. The Matrix–Tree Theorem (Theorem 5.6.8) was first proved by Borchardt [**4**] in 1860, though a similar result had earlier been published by Sylvester [**68**] in 1857. Cayley [**8**, p. 279] in fact in 1856 referred to the not yet published work of Sylvester. For further historical information on the Matrix-Tree Theorem, see [**47**, p. 42]. Typically the Matrix-Tree Theorem is proved using the Binet–Cauchy formula (a formula for the determinant of the product of an $m \times n$ matrix and an $n \times m$ matrix); see [**47**, §5.3] for such a proof. Additional information on the eigenvalues of the adjacency matrix and Laplacian matrix of a graph may be found in [**12**][**13**][**14**].

The fundamental reason underlying the simple product formula for $c(C_n)$ given by equation (5.85) is that the graph C_n has a high degree of symmetry, viz., it is a *Cayley graph* of the abelian group $\Gamma = (\mathbb{Z}/2\mathbb{Z})^n$. This is equivalent to the statement that Γ acts regularly on the vertices of C_n, i.e., Γ is transitive and only the identity

element fixes a vertex. For the complexity of an arbitrary Cayley graph of a finite abelian group, see Exercise 5.68. In general, it follows from group representation theory that the automorphism group of a graph G "induces" a factorization of the characteristic polynomial of the adjacency matrix of G; see e.g. [**13**, Ch. 5] for an exposition. For further aspects of Cayley graphs of $(\mathbb{Z}/2\mathbb{Z})^n$, see [**15**].

The de Bruijn sequences of Example 5.6.12 are named after Nicolaas Govert de Bruijn, who published his work on this subject in 1946 [**6**]. However, it was found by Stanley in 1975 that the problem of enumerating de Bruijn sequences had been posed by de Rivière [**58**] and solved by Flye Sainte-Marie in 1894 [**25**]. See [**7**] for an acknowledgment of this discovery. De Bruijn sequences have a number of interesting applications to the design of switching networks and related topics. For further information, see [**31**]. Additional references to de Bruijn sequences may be found in [**44**, p. 92].

References

5.1. T. van Aardenne-Ehrenfest and N. G. de Bruijn, Circuits and trees in oriented linear graphs, *Simon Stevin* (= *Bull. Belgian Math. Soc.*) **28** (1951), 203–217.

5.2. F. Bergeron, G. Labelle, and P. Leroux, *Combinatorial Species and Tree-like Structures*, Cambridge University Press, Cambridge, 1997.

5.3. N. Biggs, *Algebraic Graph Theory*, second edition, Cambridge University Press, Cambridge, 1993.

5.4. C. W. Borchardt, Ueber eine der Interpolation entsprechende Darstellung der Eliminations–Resultante, *J. reine angew. Math.* **57** (1860), 111–121.

5.5. T. J. I'A. Bromwich, *An Introduction to the Theory of Infinite Series*, Macmillan, London, 1931; second edition, 1941.

5.6. N. G. de Bruijn, A combinatorial problem, *Nederl. Akad. Wetensch. Proc.* **49** (1946), 758–764 (= *Indag. Math.* **8** (1946), 461–467).

5.7. N. G. de Bruijn, Acknowlegment of priority to C. Flye Sainte-Marie on the counting of circular arrangements of 2^n zeros and ones that show each n-letter word exactly once, TH-Report 75-WSK-06, Technological University Eindhoven, June, 1975.

5.8. A. Cayley, Note sur une formule pour la réversion des séries, *J. reine angew. Math.* (= *Crelle's J.*) **52** (1856), 276–284.

5.9. A. Cayley, A theorem on trees, *Quart. J. Math.* **23** (1889), 376–378; *Collected Papers*, vol. 13, Cambridge University Press, 1897, pp. 26–28.

5.10. S. Chowla, I. N. Herstein, and W. K. Moore, On recursions connected with symmetric groups, *Canad. J. Math.* **3** (1951), 328–334.

5.11. S. Chowla, I. N. Herstein, and W. R. Scott, The solutions of $x^d = 1$ in symmetric groups, *Norske Vid. Selsk. Forh. Trondheim* **25** (1952), 29–31.

5.12. F. Chung, *Spectral Graph Theory*, CBMS Regional Conference Series in Mathematics, no. 92, American Mathematical Society, Providence, 1997.

5.13. D. M. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs: Theory and Application*, Academic Press, New York, 1980; third edition, Johann Ambrosius Barth, Heidelberg/Leipzig, 1995.

5.14. D. M. Cvetković and M. Doob, Recent results in the theory of graph spectra, *Annals of Discrete Math.* **36**, Elsevier, Amsterdam/New York, 1988.

5.15. P. Diaconis and R. L. Graham, The Radon transform on $\mathbf{Z}_2^k$, *Pacific J. Math.* **118** (1985), 323–345.

5.16. A. W. M. Dress and T. Müller, Decomposable functors and the exponential principle, *Advances in Math.* **129** (1997), 188–221.
5.17. A. Dvoretzky and Th. Motzkin, A problem of arrangements, *Duke Math. J.* **14** (1947), 305–313.
5.18. N. Dershowitz and S. Zaks, The cycle lemma and some applications, *Europ. J. Combinatorics* **11** (1990), 35–40.
5.19. Editorial note, *Amer. Math. Monthly* **59** (1952), 296–297.
5.20. A. Erdélyi and I. M. H. Etherington, Some problems of non-associative combinations (2), *Edinburgh Math. Notes* **32** (1940), 7–12.
5.21. P. L. Erdős and L. A. Székely, Applications of antilexicographic order. I. An enumerative theory of trees, *Advances in Appl. Math.* **10** (1989), 488–496.
5.22. I. M. H. Etherington, Some problems of non-associative combinations (1), *Edinburgh Math. Notes* **32** (1940), 1–6.
5.23. F. Faà di Bruno, Sullo sviluppo delle funzioni, *Annali di Scienze Matematiche et Fisiche di Tortoloni* **6** (1855), 479–480.
5.24. F. Faà di Bruno, Note sur un nouvelle formule de calcul différentiel, *Quarterly J. Math.* **1** (1857), 359–360.
5.25. C. Flye Sainte-Marie, Solution to question 48, *l'Intermédiare des Mathématiciens* **1** (1894), 107–110.
5.26. D. Foata and M. P. Schützenberger, *Theorie Géometrique des Polynômes Eulériens*, Lecture Notes in Math., no. 138, Springer-Verlag, Berlin, 1970.
5.27. F. G. Frobenius, Über die Charaktere der symmetrischen Gruppe, *Sitzungsber. Kön. Preuss. Akad. Wissen. Berlin* (1900), 516–534; *Gesammelte Abh. III* (J.-P. Serre, ed.), Springer-Verlag, Berlin, 1968, pp. 148–166.
5.28. I. M. Gessel, A noncommutative generalization and q-analog of the Lagrange inversion formula, *Trans. Amer. Math. Soc.* **257** (1980), 455–482.
5.29. I. M. Gessel, A q-analog of the exponential formula, *Discrete Math.* **40** (1982), 69–80.
5.30. I. M. Gessel, A combinatorial proof of the multivariable Lagrange inversion formula, *J. Combinatorial Theory (A)* **45** (1987), 178–195.
5.31. S. W. Golomb, *Shift Register Sequences* (with portions co-authored by L. R. Welch, R. M. Goldstein, and A. W. Hales), Holden-Day, San Francisco/Cambridge/Amsterdam, 1967.
5.32. I. J. Good, Normally recurring decimals, *J. London Math. Soc.* **21** (1947), 167–169.
5.33. I. J. Good, The generalization of Lagrange's expansion and the enumeration of trees, *Proc. Camb. Phil. Soc.* **61** (1965), 499–517. Corrections, **64** (1968), 489.
5.34. R. C. Grimson, Some results on the enumeration of symmetric arrays, *Duke Math. J.* **38** (1971), 711–715.
5.35. H. Gupta, Enumeration of symmetric matrices, *Duke Math. J.* **35** (1968), 653–659.
5.36. P. Henrici, Die Lagrange-Bürmannsche Formal bein formalen Potentzreihen, *Jahresber. Deutsch. Math. Verein.* **86** (1984), 115–134.
5.37. A. Hurwitz, Über die Anzahl der Riemann'schen Flächen mit gegebenen Verweigungspunkten, *Math. Ann.* **55** (1902), 53–66.
5.38. K. G. J. Jacobi, Zur combinatorischen Analysis, *J. Reine Ang. Math.* (= *Crelle's J.*) **22** (1841), 372–374.
5.39. E. Jacobstahl, Sur le nombre d'éléments de S_n dont l'ordre est un nombre premier, *Norske Vid. Selsk. Forh. Trondheim* **21** (1949), 49–51.
5.40. D. E. Knuth, Oriented subtrees of an arc digraph, *J. Combinatorial Theory* **3** (1967), 309–314.
5.41. D. E. Knuth, *The Art of Computer Programming*, vol. 1, *Fundamental Algorithms*, Addison-Wesley, Reading, Massachusetts, 1968; second edition, 1973.
5.42. G. Labelle, Une nouvelle démonstration combinatoire des formules d'inversion de Lagrange, *Advances in Math.* **42** (1981), 217–247.

5.43. L. de Lagrange, Nouvelle méthode pour résoudre des équations littérales par le moyen des séries, *Mém. Acad. Roy. Sci. Belles-Lettres de Berlin* **24** (1770).

5.44. J. H. van Lint, *Combinatorial Theory Seminar, Eindhoven University of Technology*, Lecture Notes in Mathematics, no. 382, Springer-Verlag, Berlin/Heidelberg/New York, 1974.

5.45. J. W. Moon, The second moment of the complexity of a graph, *Mathematika* **11** (1964), 95–98.

5.46. J. W. Moon, Various proofs of Cayley's formula for counting trees, in *A Seminar on Graph Theory* (F. Harary, ed.), Holt, Rinehart, and Winston, New York, 1967, pp. 70–78.

5.47. J. W. Moon, *Counting Labelled Trees*, Canadian Math. Monographs, no. 1, Canadian Mathematical Congress, 1970.

5.48. N. Metropolis and G.-C. Rota, Witt vectors and the algebra of necklaces, *Advances in Math.* **50** (1983), 95–125.

5.49. E. H. Neville, The codifying of tree structure, *Proc. Camb. Phil. Soc.* **49** (1953), 381–385.

5.50. G. Pólya, Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und Chemische Verbindungen, *Acta Math.* **68** (1937), 145–245; reprinted in *Collected Papers*, vol. 4, M.I.T. Press, Cambridge, Massachusetts, 1984, pp. 308–416.

5.51. G. Pólya, Partitions of a finite set into structured subsets, *Math. Proc. Camb. Phil. Soc.* **77** (1975), 453–458; reprinted in *Collected Papers*, vol. 4, M.I.T. Press, Cambridge, Massachusetts, 1984, pp. 464–469.

5.52. H. Prüfer, Neuer Beweis eines Satzes über Permutationen, *Arch. Math. Phys.* **27** (1918), 142–144.

5.53. G. Pólya and G. Szegö, *Problems and Theorems in Analysis I, II*, Springer-Verlag, New York/Heidelberg/Berlin, 1972, 1976.

5.54. G. N. Raney, Functional composition patterns and power series reversion, *Trans. Amer. Math. Soc.* **94** (1960), 441–451.

5.55. G. N. Raney, A formal solution of $\sum_{i=1}^{\infty} A_i e^{B_i X} = X$, *Canad. J. Math.* **16** (1964), 755–762.

5.56. R. J. Riddell and G. E. Uhlenbeck, On the theory of the virial development of the equation of state of monoatomic gases, *J. Chem. Phys.* **21** (1953), 2056–2064.

5.57. J. Riordan, The enumeration of labelled trees by degrees, *Bull. Amer. Math. Soc.* **72** (1966), 110–112.

5.58. A. de Rivière, Question 48, *l'Intermédiare des Mathématiciens* **1** (1894), 19–20.

5.59. J. Rosen, The number of product-weighted lead codes for ballots and its relation to the Ursell functions of the linear Ising model, *J. Combinatorial Theory (A)* **20** (1976), 377–384.

5.60. E. Schröder, Vier combinatorische Probleme, *Z. für Math. Phys.* **15** (1870), 361–376.

5.61. I. Schur, problem, *Arch. Math. Phys. Series 3* **27** (1918), 162.

5.62. M. P. Schützenberger, Le théorème de Lagrange selon G. N. Raney, *Sém. IRIA, Logiques et Automates*, 1971, pp. 199–205.

5.63. D. Singer, Q-analogues of Lagrange inversion, *Advances in Math.* **115** (1995), 99–116.

5.64. C. A. B. Smith and W. T. Tutte, On unicursal paths in a network of degree 4, *Amer. Math. Monthly* **48** (1941), 233–237.

5.65. R. Stanley, Exponential structures, *Studies in Applied Math.* **59** (1978), 73–82.

5.66. D. W. Stanton, Recent results for the q-Lagrange inversion formula, in *Ramanujan Revisted* (G. E. Andrews *et al.*, eds.), Academic Press, San Diego/London, 1988, pp. 525–536.

5.67. G. S. Sylvester, *Continuous-Spin Ising Ferromagnets*, Ph.D. thesis, Massachusetts Institute of Technology, 1976.
5.68. J. J. Sylvester, On the change of systems of independent variables, *Quart. J. Math.* **1** (1857), 42–56; *Collected Mathematical Papers*, vol. 2, Cambridge, 1908, pp. 65–85.
5.69. J. Touchard, Sur les cycles des substitutions, *Acta Math.* **70** (1939), 243–297.
5.70. W. T. Tutte, The dissection of equilateral triangles into equilateral triangles, *Proc. Cambridge Phil. Soc.* **44** (1948), 463–482.

Exercises

5.1. a. [2–] Each of n (distinguishable) telephone poles is painted red, white, blue, or yellow. An odd number are painted blue and an even number yellow. In how many ways can this be done?

b. [2] Suppose now the colors orange and purple are also used. The number of orange poles plus the number of purple poles is even. Now how many ways are there?

5.2. a. [3–] Write

$$1+\sum_{n\geq 1} f_n x^n = \exp \sum_{n\geq 1} h_n \frac{x^n}{n},$$

where $h_n \in \mathbb{Q}$ (or any field of characteristic 0). Show that the following four conditions are equivalent for fixed $N \in \mathbb{P}$:

(i) $f_n \in \mathbb{Z}$ for all $n \in [N]$.

(ii) $h_n \in \mathbb{Z}$ and $\sum_{d|n} h_d\,\mu(n/d) \equiv 0 \pmod n$ for all $n \in [N]$, where μ denotes the ordinary number-theoretic Möbius function.

(iii) $h_n \in \mathbb{Z}$ for all $n \in [N]$, and $h_n \equiv h_{n/p} \pmod{p^r}$, whenever $n \in [N]$ and p is a prime such that $p^r \mid n,\ p^{r+1} \nmid n,\ r \geq 1$.

(iv) There exists a polynomial $P(t) = \prod_1^N (t-\alpha_i) \in \mathbb{Z}[t]$ (where $\alpha_i \in \mathbb{C}$) such that $h_n = \sum_{i=1}^N \alpha_i^n$ for all $n \in [N]$.

b. [2+] (basic knowledge of finite fields required) Let S be a set of polynomial equations in the variables $x_1, \ldots, x_k$ over the field $\mathbb{F}_q$. Let N_n denote the number of solutions $(\alpha_1, \ldots, \alpha_k)$ to the equations such that each $\alpha_i \in \mathbb{F}_{q^n}$. Show that the generating function

$$Z(x) = \exp \sum_{n\geq 1} N_n \frac{x^n}{n}$$

has integer coefficients.

c. [3–] Show that if $\alpha_1, \ldots, \alpha_N \in \mathbb{C}$ and $\sum_{i=1}^N \alpha_i^n \in \mathbb{Z}$ for all $n \in \mathbb{P}$, then $\prod_{i=1}^N (t-\alpha_i) \in \mathbb{Z}[t]$.

5.3. a. [2–] Let $f(n) = 1\cdot 3\cdot 5\cdots(2n-1)$ and $g(n) = 2^n n!$. Show that $E_g(x) = E_f(x)^2$.

b. [3–] Give a combinatorial proof based on Proposition 5.1.1.

5.4. a. [2] A *threshold graph* is a simple (i.e., no loops or multiple edges) graph which may be defined inductively as follows:

(i) The empty graph is a threshold graph.

(ii) If G is a threshold graph, then so is the disjoint union of G with a one-vertex graph.

(iii) If G is a threshold graph, then so is the (edge) complement of G.

Let $t(n)$ denote the number of threshold graphs with vertex set $[n]$ (with $t(0)=1$), and let $s(n)$ denote the number of such graphs with no isolated vertex (so $s(0)=1, s(1)=0$). Set

$$T(x) = E_t(x) = 1 + x + 2\frac{x^2}{2!} + 8\frac{x^3}{3!} + 46\frac{x^4}{4!} + \cdots,$$

$$S(x) = E_s(x) = 1 + \frac{x^2}{2!} + 4\frac{x^3}{3!} + 23\frac{x^4}{4!} + \cdots.$$

Show that $T(x)=e^x S(x)$ and $T(x)=2S(x)+x-1$ to deduce

$$T(x) = e^x(1-x)/(2-e^x),$$

$$S(x) = (1-x)/(2-e^x). \tag{5.86}$$

b. [2] Let $c(n)$ denote the number of ordered partitions (or preferential arrangements) of $[n]$, so by Example 3.15.10 $E_c(x)=1/(2-e^x)$. It follows from (5.86) that $s(n)=c(n)-nc(n-1)$. Give a direct combinatorial proof.

c. [3–] Let $\mathcal{T}_n$ denote the set of all hyperplanes $x_i + x_j = 0$, $1 \le i < j \le n$, in $\mathbb{R}^n$. The hyperplane arrangement $\mathcal{T}_n$ is called the *threshold arrangement.* Show that the number of regions of $\mathcal{T}_n$ (i.e., the number of connected components of the space $\mathbb{R}^n - \bigcup_{H\in\mathcal{T}_n} H$) is equal to $t(n)$.

d. [3–] Let L_n be the intersection poset of $\mathcal{T}_n$, as defined in Exercise 3.56. Show that the characteristic polynomial of L_n is given by

$$\sum_{n\ge 0}(-1)^n\chi(L_n,-q)\frac{x^n}{n!} = (1-x)\left(\frac{e^x}{2-e^x}\right)^{\frac{q+1}{2}}.$$

This result generalizes (c), since by Exercise 3.56(a) the number of regions of $\mathcal{S}_n$ is equal to $|\chi(L_n,-1)|$.

5.5. [2+] Let $b_k(n)$ be the number of bipartite graphs (without multiple edges) with k edges on the vertex set $[n]$. For instance, $b_0(3)=1$, $b_1(3)=3$, $b_2(3)=3$, and $b_3(3)=0$. Show that

$$\sum_{n\ge 0}\sum_{k\ge 0} b_k(n)q^k\frac{x^n}{n!} = \left[\sum_{n\ge 0}\left(\sum_{i=0}^{n}(1+q)^{i(n-i)}\binom{n}{i}\right)\frac{x^n}{n!}\right]^{1/2}.$$

5.6. [2] Let $\chi(K_{mn},q)$ denote the chromatic polynomial (as defined in Exercise 3.44) of the complete bipartite graph K_{mn}. Show that

$$\sum_{m,n\ge 0}\chi(K_{mn},q)\frac{x^m}{m!}\frac{y^n}{n!} = (e^x+e^y-1)^q.$$

5.7. In this exercise we develop the rudiments of the theory of "combinatorial trigonometry." Let E_n be the number of alternating permutations π of $[n]$, as discussed

at the end of Chapter 3.16. Thus $\pi = a_1 a_2 \cdots a_n$, where $a_1 > a_2 < a_3 > \cdots a_n$.

a. [2] Using the fact (equation (3.58)) that

$$\sum_{n\geq 0} E_n \frac{x^n}{n!} = \tan x + \sec x,$$

give a *combinatorial* proof that $1 + \tan^2 x = \sec^2 x$.

b. [2+] Do the same for the identity

$$\tan(x+y) = \frac{\tan x + \tan y}{1 - (\tan x)(\tan y)}. \tag{5.87}$$

* **5.8. a.** [2] The *central factorial numbers* $T(n,k)$ are defined for $n, k \in \mathbb{N}$ by

$$T(n,0) = T(0,k) = 0, \qquad T(1,1) = 1,$$
$$T(n,k) = k^2 T(n-1,k) + T(n-1,k-1) \qquad \text{for } (n,k) \in \mathbb{P}^2 - \{(1,1)\},$$

Show that

$$T(n,k) = 2\sum_{j=1}^{k} \frac{j^{2n}(-1)^{k-j}}{(k-j)!\,(k+j)!}$$

and

$$\sum_{n\geq 0} T(n,k)\frac{x^{2n}}{(2n)!} = \frac{1}{(2k)!}\left(2\sinh\frac{x}{2}\right)^{2k}. \tag{5.88}$$

b. [2] Show that

$$\sum_{n\geq 0} T(n,k)x^n = \frac{x^k}{(1-1^2x)(1-2^2x)\cdots(1-k^2x)}.$$

c. [2] Show that $T(n,k)$ is equal to the number of partitions of the set $\{1, 1', 2, 2', \ldots, n, n'\}$ into k blocks, such that for every block B, if i is the least integer for which $i \in B$ or $i' \in B$, then both $i \in B$ and $i' \in B$.

d. [2+] The *Genocchi numbers* G_n are defined by

$$\frac{2x}{e^x+1} = \sum_{n\geq 1} G_n \frac{x^n}{n!}$$
$$= x - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{3x^6}{6!} + \frac{17x^8}{8!} - \frac{155x^{10}}{10!} + \frac{2073x^{12}}{12!} - \cdots.$$

Show that $G_{2n+1} = 0$ if $n \geq 1$, and that $(-1)^n G_{2n}$ is an odd positive integer. (Sometimes $(-1)^n G_{2n}$ is called a Genocchi number.) Note also that

$$x\tan\frac{x}{2} = \sum_{n\geq 1} (-1)^n G_{2n} \frac{x^{2n}}{(2n)!}.$$

e. [3] Show that

$$G_{2n+2} = \sum_{i=1}^{n} (-1)^{i+1} (i!)^2 T(n, i).$$

f. [3] Show that $(-1)^n G_{2n}$ counts the following:

(i) The number of permutations $\pi \in \mathfrak{S}_{2n-2}$ such that $1 \leq \pi(2i-1) \leq 2n-2i$ and $2n-2i \leq \pi(2i) \leq 2n-2$.

(ii) The number of permutations $\pi \in \mathfrak{S}_{2n-1}$ with descents after even numbers and ascents after odd numbers, e.g., 2143657 and 3564217. (Such permutations must end with $2n-1$.)

(iii) The number of pairs $(a_1, a_2, \dots, a_{n-1})$ and $(b_1, b_2, \dots, b_{n-1})$ such that $a_i, b_i \in [i]$ and every $j \in [n-1]$ occurs at least once among the a_i's and b_i's.

(iv) The number of reverse alternating permutations $a_1 < a_2 > a_3 < a_4 > \cdots > a_{2n-1}$ of $[2n-1]$ whose inversion table (as defined in Section 1.3) has only even entries. For example, for $n = 3$ we have the three permutations 45231, 34251, 24153 with inversion tables 42200, 42000, 20200.

5.9. Let $\mathcal{S}$ be a "structure" that can be put on a finite set by choosing a partition of S and putting a "connected" structure on each block, so that the exponential formula (Corollary 5.1.6) is applicable. Let $f(n)$ be the number of structures that can be put on an n-set, and let $F(x) = E_f(x)$, the exponential generating function of f.

a. [2–] Let $g(n)$ be the number of structures that can be put on an n-set so that every connected component has even cardinality. Show that

$$E_g(x) = \sqrt{F(x)F(-x)}.$$

b. [2] Let $e(n)$ be the number of structures that can be put on an n-set so that the number of connected components is even. Show that

$$E_e(x) = \frac{1}{2}\left(F(x) + \frac{1}{F(x)}\right).$$

5.10. a. [2–] Let $k > 2$. Give a generating function proof that the number $f_k(n)$ of permutations $\pi \in \mathfrak{S}_n$ all of whose cycle lengths are divisible by k is given by

$$1^2 \cdot 2 \cdot 3 \cdots (k-1)(k+1)^2(k+2) \cdots (2k-1)(2k+1)^2(2k+2) \cdots (n-1)$$

if $k \mid n$, and is 0 otherwise.

b. [2] Give a combinatorial proof of (a).

c. [2] Let $k \in \mathbb{P}$. Give a generating-function proof that the number $g_k(n)$ of permutations $\pi \in \mathfrak{S}_n$ none of whose cycle lengths is divisible by k is given by

$$1 \cdot 2 \cdots (k-1)^2(k+1) \cdots (2k-2)(2k-1)^2(2k+1) \cdots (n-1)n \qquad \text{if } k \nmid n$$

$$1 \cdot 2 \cdots (k-1)^2(k+1) \cdots (2k-2)(2k-1)^2(2k+1) \cdots (n-2)(n-1)^2 \qquad \text{if } k \mid n.$$

d. [3–] Give a combinatorial proof of (c).

5.11. **a.** [2] Let $a(n)$ be the number of permutations w in $\mathfrak{S}_n$ that have a square root, i.e., there exists $u \in \mathfrak{S}_n$ satisfying $u^2 = w$. Show that

$$\sum_{n\geq 0} a(n)\frac{x^n}{n!} = \left(\frac{1+x}{1-x}\right)^{1/2} \prod_{k\geq 1} \cosh \frac{x^{2k}}{2k}.$$

b. [2–] Deduce from (a) that $a(2n+1) = (2n+1)a(2n)$. Is there a simple combinatorial proof?

5.12. [2+] Let $f(n)$ be the number of pairs (u, v) of permutations in $\mathfrak{S}_n$ satisfying $u^2 = v^2$. Find the exponential generating function $F(x) = \sum_{n\geq 0} f(n)x^n/n!$.

5.13. **a.** [2+] Let G be a finitely generated group, and let $\mathrm{Hom}(G, \mathfrak{S}_n)$ denote the set of homomorphisms $G \to \mathfrak{S}_n$. Let $j_d(G)$ denote the number of subgroups of G of index d. Show that

$$\sum_{n\geq 0} \#\mathrm{Hom}(G, \mathfrak{S}_n)\frac{x^n}{n!} = \exp\left(\sum_{d\geq 1} j_d(G)\frac{x^d}{d}\right).$$

Note that equation (5.31) is equivalent to the case $G = \mathbb{Z}/r\mathbb{Z}$.

b. [1+] Let F_s denote the free group on s generators. Deduce from (a) that

$$\sum_{n\geq 0} n!^{s-1}x^n = \exp\left(\sum_{d\geq 1} j_d(F_s)\frac{x^d}{d}\right). \tag{5.89}$$

c. [3–] With G as above, let $u_d(G)$ denote the number of conjugacy classes of subgroups of G of index d. In particular, if every subgroup of G of index d is normal (e.g., if G is abelian) then $u_d(G) = j_d(G)$. Show that

$$\sum_{n\geq 0} \#\mathrm{Hom}(G \times \mathbb{Z}, \mathfrak{S}_n)\frac{x^n}{n!} = \prod_{d\geq 1}(1 - x^d)^{-u_d(G)}. \tag{5.90}$$

d. [1+] Let $c_m(n)$ be the number of commuting m-tuples $(u_1, \dots, u_m) \in \mathfrak{S}_n^m$, i.e., $u_i u_j = u_j u_i$ for all i and j. Deduce from (c) that

$$\sum_{n\geq 0} c_m(n)\frac{x^n}{n!} = \prod_{d\geq 1}(1 - x^d)^{-j_d(\mathbb{Z}^{m-1})}.$$

e. [3–] Let $h_k(n)$ be the number of graphs (with multiple edges allowed) on the vertex set $[n]$ with edges colored $1, 2, \dots, k-1$ satisfying the following properties:

(i) For each i, the edges colored i have no vertices in common.

(ii) For each $i < k-1$, every connected component of the (spanning) subgraph consisting of all edges colored i and $i+1$ is either a single vertex, a path of length two, a two-cycle (that is, an edge colored i and an edge colored $i+1$ with the same vertices), or a six-cycle.

(iii) For each i, j such that $j - i \geq 2$, every connected component of the subgraph consisting of all edges colored i and j is either a single vertex, a single edge (colored either i or j), a two-cycle, or a four-cycle.

Show that

$$\sum_{n\geq 0} h_k(n)\frac{x^n}{n!} = \exp\left(\sum_{d|k!} j_d(\mathfrak{S}_k)\frac{x^d}{d}\right).$$

5.14. **a.** [2–] Let $A_n(t)$ denote an Eulerian polynomial, as defined in Section 1.3, and set $y = \sum_{n\geq 1} A_n(t)x^n/n!$. Show that y is the unique power series for which there exists a power series z satisfying the two formulas

$$1 + y = \exp(tx + z)$$
$$1 + t^{-1}y = \exp(x + z).$$

b. [2+] Show that the power series z of (a) is given by

$$z = \sum_{n\geq 2} A_{n-1}(t)\frac{x^n}{n!}.$$

c. [2+] Set $(1 + y)^q = \sum_{n\geq 0} B_n(q, t)x^n/n!$. Show that

$$B_n(q, t) = \sum_{w\in\mathfrak{S}_n} q^{m(w)}t^{1+d(w)},$$

where $m(w)$ denotes the number of left-to-right minima of w, and $d(w)$ denotes the number of descents of w.

d. [2–] Deduce that the coefficient of $x^n/n!$ in $(1 + y)^{q/t}$ is a polynomial in q and t with integer coefficients.

5.15. [2] For each of the following sets of graphs, let $f(n)$ be the number of graphs G on the vertex set $[n]$ such that every connected component of G is isomorphic to some graph in the set. Find for each set $E_f(x) = \sum_{n\geq 0} f(n)x^n/n!$. (Set $f(0) = 1$.)

a. Cycles C_i of length $i \geq k$ (for some fixed $k \geq 3$)

b. Stars K_{1i}, $i \geq 1$ (K_{rs} denotes a complete bipartite graph)

c. Wheels W_i with $i \geq 4$ vertices (W_i is obtained from C_{i-1} by adding a new vertex joined to every vertex of C_{i-1})

d. Paths P_i with $i \geq 1$ vertices (so P_1 is a single vertex and P_2 is a single edge).

5.16. Let G be a simple graph (i.e., no loops or multiple edges) on the vertex set $[n]$. The (ordered) *degree sequence* of G is defined to be $d(G) = (d_1, \ldots, d_n)$, where d_i is the degree (number of incident edges) of vertex i. Let $f(n)$ be the number of *distinct* degree sequences of simple graphs on the vertex set $[n]$. For instance, all eight graphs on [3] have different degree sequences, so $f(3) = 8$. On the other hand, there are three graphs on [4] with degree sequence (1, 1, 1, 1), so $f(4) < 2^{\binom{4}{2}} = 64$. (In fact, $f(4) = 54$.)

a. [3+] Show that

$$f(n) = \sum_X \max\{1, 2^{c(X)-1}\}, \tag{5.91}$$

where X ranges over all graphs on $[n]$ such that every connected component is either a tree or has a single cycle, and all cycles of X are of odd length; and where $c(X)$ denotes the number of (odd) cycles of X.

b. [3–] Let

$$F(x) = \sum_{n\geq 0} f(n)\frac{x^n}{n!}$$

$$= 1 + x + 2\frac{x^2}{2!} + 8\frac{x^3}{3!} + 54\frac{x^4}{4!} + 533\frac{x^5}{5!} + 6944\frac{x^6}{6!} + \cdots.$$

Assuming (a), show that

$$F(x) = \frac{1}{2}\left[\left(1 + 2\sum_{n\geq 1} n^n \frac{x^n}{n!}\right)^{1/2}\left(1 - \sum_{n\geq 1}(n-1)^{n-1}\frac{x^n}{n!}\right) + 1\right]$$
$$\times e^{\sum_{n\geq 1} n^{n-2}x^n/n!},$$

where we set $0^0 = 1$ in the term $n = 1$ of the second sum on the right.

5.17. a. [2] Fix $k, n \in \mathbb{P}$. In how many ways may n people form exactly k lines? (In other words, how many ways are there of partitioning the set $[n]$ into k blocks, and then linearly ordering each block?) Give a simple combinatorial proof.

b. [2–] Deduce that

$$1 + \sum_{n\geq 1}\sum_{k=1}^{n} \frac{n!}{k!}\binom{n-1}{k-1} x^k \frac{u^n}{n!} = \exp\frac{xu}{1-u}.$$

c. [2+] Let $a \in \mathbb{P}$. Extend the argument of (a) to deduce that

$$1 + \sum_{n\geq 1}\sum_{k=1}^{n} \frac{n!}{k!}\binom{n+(a-1)k-1}{n-k} x^k \frac{u^n}{n!} = \exp\frac{xu}{(1-u)^a} \qquad (5.92)$$

and

$$1 + \sum_{n\geq 1}\sum_{k=1}^{n} \frac{n!}{k!}\binom{ak}{n-k} x^k \frac{u^n}{n!} = \exp xu(1+u)^a. \qquad (5.93)$$

Note. Since these identities hold for all $a \in \mathbb{P}$, they also hold if a is an indeterminate.

d. [2] Fix $k, n, \alpha \in \mathbb{N}$. Let A be a set of cardinality α disjoint from $[n]$. In how many ways can we choose a subset S of $[n]$, then choose a partition π of S into exactly k blocks, then linearly order each block of π, and finally choose an injection $f : \bar{S} \to \bar{S} \cup A$, where $\bar{S} = [n] - S$? Give a simple combinatorial proof.

e. [2–] Deduce that

$$\sum_{n\geq 0}\sum_{k=0}^{n}\binom{n}{k}(\alpha+n)_{n-k}\, x^k \frac{u^n}{n!} = (1-u)^{-\alpha-1}\exp\frac{xu}{1-u}.$$

(Note that we obtain (b) by setting $\alpha = -1$.)

5.18. [2] Call two permutations $\pi, \sigma \in \mathfrak{S}_n$ *equivalent* if every cycle C of π is a power D^j (where j depends on C) of some cycle D of σ. Clearly this is an equivalence relation; let $e(n)$ be the number of equivalence classes (with $e(0) = 1$). Show that

$$\sum_{n\geq 0} e(n)\frac{x^n}{n!} = \exp \sum_{n\geq 1} \frac{x^n}{n\phi(n)},$$

where ϕ is Euler's phi-function.

5.19. [3–] Define polynomials $K_n(a)$ by

$$\sum_{n\geq 0} K_n(a)\frac{u^n}{n!} = \exp\left(au + \frac{u^2}{2}\right).$$

Thus it follows from Example 5.2.10 that

$$K_n(a) = \sum_{\pi} a^{c_1(\pi)}, \tag{5.94}$$

where π ranges over all involutions (i.e., $\pi^2 = 1$) in $\mathfrak{S}_n$, and $c_1(\pi)$ is the number of 1-cycles (fixed points) of π. Using (5.94), give a combinatorial proof of the identity

$$\sum_{n\geq 0} K_n(a)K_n(b)\frac{x^n}{n!} = (1 - x^2)^{-1/2} \exp\left[\frac{abx + \frac{1}{2}(a^2 + b^2)x^2}{1 - x^2}\right]. \tag{5.95}$$

5.20. a. [2+] A *block* is a finite connected graph B (allowing multiple edges but not loops) with at least two vertices such that the removal of any vertex v and all edges incident to v leaves a connected graph. Let $\mathcal{B}$ be a collection of nonisomorphic blocks. Let $b(n)$ be the number of blocks on the vertex set $[n]$ which are isomorphic to some block in $\mathcal{B}$. In other words, if Aut B denotes the automorphism group of the block B, then

$$b(n) = \sum_{B} \frac{n!}{\#(\text{Aut } B)},$$

summed over all n-vertex blocks B in $\mathcal{B}$. Call a graph G a *$\mathcal{B}$-graph* if it is connected and its maximal blocks (i.e., maximal induced subgraphs which are blocks) are all isomorphic to members of $\mathcal{B}$. For $n \geq 2$, let $f(n)$ be the number of *rooted* $\mathcal{B}$-graphs on an n-element vertex set V (i.e., a $\mathcal{B}$-graph with a vertex chosen as a root). Set $f(0) = 0$ and $f(1) = 1$, and put

$$B(x) = E_b(x) = \sum_{n\geq 2} b(n)\frac{x^n}{n!}$$

$$F(x) = E_f(x) = \sum_{n\geq 1} f(n)\frac{x^n}{n!}.$$

Show that

$$F(x) = xe^{B'(F(x))}, \tag{5.96}$$

and hence

$$\sum_{n\geq 1} b(n+1)\frac{x^n}{n!} = \log\left(\frac{x}{F^{\langle -1\rangle}(x)}\right). \tag{5.97}$$

For instance, if $\mathcal{B}$ contains only the single block consisting of one edge, then a $\mathcal{B}$-graph is a (free) tree. Hence $f(n)$ is the number of rooted trees on n vertices, $B(x)=x^2/2!$, and $F(x)=xe^{F(x)}$ (agreeing with Proposition 5.3.1).

b. [2] Let $g(n)$ be the total number of blocks without multiple edges on an n-element vertex set. Show that

$$\sum_{n\geq 1} g(n+1)\frac{x^n}{n!} = \log\left(\frac{x}{G(x)^{\langle -1\rangle}}\right),$$

where

$$G(x) = \frac{\sum_{n\geq 1} 2^{\binom{n}{2}}\frac{x^n}{(n-1)!}}{\sum_{n\geq 0} 2^{\binom{n}{2}}\frac{x^n}{n!}}.$$

5.21. [3–] Find a combinatorial proof of equation (4.39). More specifically, using the notation of Chapter 4.7, given a pair (π, u), where $\pi \in \mathfrak{S}_n$ and $u \in \mathcal{B}_n^*$, associate with it in bijective fashion a permutation $\sigma \in \mathfrak{S}_n$ with a cyclic shift v_C of an element of $\mathcal{B}_k^*$ attached to each k-cycle C of σ. The multiset of letters in u should coincide with those in all the v_C's so that the bijection is weight-preserving.

5.22. [2] Let $L(n)$ be the function of Example 5.2.9, so in particular $L(n)$ is the number of graphs on the vertex set $[n]$ for which every component is a cycle (including loops and double edges). Give a direct combinatorial proof that

$$L(n+1) = (n+1)L(n) - \binom{n}{2}L(n-2), \qquad n \geq 2.$$

5.23. [2] Let Δ be a set $\{\delta_1, \ldots, \delta_n\}$ of n straight lines in the plane lying in general position (i.e., no two are parallel and no three meet at a point). Let P be the set of their points $\delta_i \cap \delta_j$ of intersection, so $\#P = \binom{n}{2}$. A *cloud* is an n-subset of P containing no three collinear points. Find a bijection between clouds and regular graphs on $[n]$ (without loops and multiple edges) of degree two. Hence, by (5.29), if $c(n)$ is the number of clouds for $\#\Delta = n$, then

$$\sum_{n\geq 0} c(n)\frac{x^n}{n!} = (1-x)^{-1/2}\exp\left(-\frac{x}{2} - \frac{x^2}{4}\right).$$

5.24. **a.** [2+] Let Σ_n be the convex polytope of all $n \times n$ *symmetric* doubly stochastic matrices. Show that the extreme points (vertices) of Σ_n consist of all matrices $\frac{1}{2}(P + P^t)$, where P is a permutation matrix corresponding to a permutation with no cycles of even length ≥ 4.

b. [2+] Let $M(n)$ be the number of vertices of Σ_n. Show that

$$\sum_{n\geq 0} M(n)\frac{x^n}{n!} = \left(\frac{1+x}{1-x}\right)^{1/4}\exp\left(\frac{x}{2} + \frac{x^2}{2}\right). \tag{5.98}$$

c. [2+] Find polynomials $p_0(n), \ldots, p_3(n)$ such that

$$M(n+1) = p_0(n)M(n)+p_1(n)M(n-1)+p_2(n)M(n-2)+p_3(n)M(n-3),$$

for all $n \geq 3$.

* d. [5–] Is there a direct combinatorial proof of (c), analogous to Exercise 5.22?

5.25. a. [2+] Let Σ_n^* be the convex polytope of all $n \times n$ *symmetric substochastic* matrices (i.e., the entries are ≥ 0, and all line sums are ≤ 1). Show that the vertices of Σ_n^* are obtained from those of Σ_n (defined in the previous exercise) by replacing some 1's on the main diagonal by 0's.

b. [2] Let $M^*(n)$ be the number of vertices of Σ_n^*. Show that

$$\sum_{n\geq 0} M^*(n)\frac{x^n}{n!} = e^x \sum_{n\geq 0} M(n)\frac{x^n}{n!},$$

where $M(n)$ is defined in Exercise 5.24.

c. [2] Find polynomials $p_0^*(n), \ldots, p_3^*(n)$ such that

$$\begin{aligned} M^*(n+1) = p_0^*(n)M^*(n) &+ p_1^*(n)M^*(n-1) + p_2^*(n)M^*(n-2) \\ &+ p_3^*(n)M^*(n-3). \end{aligned}$$

5.26. [2+] Let $f(n)$ be the number of sets S of nonempty subsets of $[n]$ (including $S = \emptyset$) such that any two elements of S are either disjoint or comparable (with respect to inclusion). Let $g(n)$ be the number of such sets S which contain $[n]$, with $g(0) = 0$. Set

$$F(x) = E_f(x) = 1 + 2x + 8\frac{x^2}{2!} + 64\frac{x^3}{3!} + 832\frac{x^4}{4!} + 15104\frac{x^5}{5!} + \cdots$$

$$G(x) = E_g(x) = \qquad x + 4\frac{x^2}{2!} + 32\frac{x^3}{3!} + 416\frac{x^4}{4!} + 7552\frac{x^5}{5!} + \cdots.$$

Show that $F(x) = 1 + 2G(x)$ and $F(x) = e^{x+G(x)}$. Hence [why?]

$$G(x) = (\log(1+2x) - x)^{\langle -1\rangle} \tag{5.99}$$

$$F(x) - 1 = \left(\log(1+x) - \frac{x}{2}\right)^{\langle -1\rangle}.$$

5.27. [2] Find the number $e(n)$ of trees with $n+1$ unlabeled vertices and n labeled *edges*. Give a simple bijective proof.

5.28. [2+] Let $k \in \mathbb{P}$. A *k-edge colored tree* is a tree whose edges are colored from a set of k colors such that any two edges with a common vertex have different colors. Show that the number $T_k(n)$ of k-edge colored trees on the vertex set $[n]$ is given by

$$T_k(n) = k(nk-n)(nk-n-1)\cdots(nk-2n+3) = k(n-2)!\binom{nk-n}{n-2}.$$

5.29. a. [2] Let P_n be the set of all planted forests on $[n]$. Let uv be an edge of a forest $F \in P_n$ such that u is closer than v to the root of its component. Define F to

cover the rooted forest F' if F' is obtained by removing the edge uv from F, and rooting the new tree containing v at v. This definition of cover defines the covering relation of a partial order on P_n. Under this partial order P_n is graded of rank $n-1$. The rank of a forest F in P_n is its number of edges. Show that an element F of P_n of rank i covers i elements and is covered by $(n-i-1)n$ elements.

b. [2] By counting in two ways the number of maximal chains of P_n, deduce that the number $r(n)$ of rooted trees on $[n]$ is equal to n^{n-1}.

c. [2+] Let $\bar{P}_n$ be P_n with a $\hat{1}$ adjoined. Show that

$$\mu(\hat{0}, \hat{1}) = (-1)^n (n-1)^{n-1},$$

where μ denotes the Möbius function of $\bar{P}_n$.

5.30. [2+] Let $R = \{1, 2, \ldots, r\}$ and $S = \{1', 2', \ldots, s'\}$ be disjoint sets of cardinalities r and s, respectively. A *free bipartite tree* with vertex bipartition (R, S) is a free tree T on the vertex set $R \cup S$ such that every edge of T is incident to a vertex in R and a vertex in S. By modifying the two proofs of Theorem 5.3.4, give two combinatorial proofs that

$$\sum_T \left(\prod_{i\in R} x_i^{\deg i}\right)\left(\prod_{j'\in S} y_j^{\deg j'}\right)$$
$$= (x_1 \cdots x_r)(y_1 \cdots y_s)(x_1 + \cdots + x_r)^{s-1}(y_1 + \cdots + y_s)^{r-1}, \qquad (5.100)$$

summed over all free bipartite trees T with vertex bipartition (R, S). In particular, the total number of such trees (i.e., the *complexity* $c(K_{rs})$ of the complete bipartite graph K_{rs}) is $r^{s-1}s^{r-1}$, agreeing with the computation at the end of the solution to Exercise 2.11(b).

5.31. **a.** [1+] Let S and T be finite sets, and for each $t \in T$ let x_t be an indeterminate. Show that

$$\sum_{f:S\to T} \prod_{s\in S} x_{f(s)} = \left(\sum_{t\in T} x_t\right)^{\#S},$$

where the first sum ranges over all functions $f: S \to T$.

b. [3–] By considering the case $S = [n]$ and $T = [n+2]$, show that

$$(x_1+\cdots+x_{n+2})^n = \sum_{A\subseteq[n]} x_{n+1}\left(x_{n+1} + \sum_{i\in A} x_i\right)^{\#A-1}\left(x_{n+2} + \sum_{i\in A'} x_i\right)^{n-\#A}.$$

where $A' = [n] - A$. Note that when $A = \emptyset$, we have

$$x_{n+1}\left(x_{n+1} + \sum_{i\in A} x_i\right)^{\#A-1} = 1.$$

c. [2–] Deduce from (b) that

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x(x-kz)^{k-1}(y+kz)^{n-k},$$

where x, y, z are indeterminates. Note that the case $z=0$ is the binomial theorem.

d. [2–] Deduce from (c) the identity

$$\sum_{n\geq 0}(n+1)^n\frac{x^n}{n!}=\left(\sum_{n\geq 0}n^n\frac{x^n}{n!}\right)\left(\sum_{n\geq 0}(n+1)^{n-1}\frac{x^n}{n!}\right),$$

where we set $0^0=1$.

5.32. a. [2+] Let $f:[n]\to[n]$, and let D_f denote the digraph of f, i.e., the directed graph on the vertex set $[n]$ with an arrow from i to j if $f(i)=j$. Thus every connected component of D_f contains a unique cycle, and every vertex i of this cycle is the root of a rooted tree (possibly consisting of the single point i) directed toward i. Let $w_f(i)=t_{jk}$ (an indeterminate) if vertex i is at distance k from a j-cycle of D_f. Let

$$w(f)=\prod_{i=1}^{n}w_f(i).$$

For instance, if D_f is given by Figure 5-18, then $w_f(1)=t_{31}$, $w_f(2)=t_{30}$, $w_f(3)=t_{30}$, $w_f(4)=t_{11}$, $w_f(5)=t_{31}$, $w_f(6)=t_{10}$, $w_f(7)=t_{32}$, $w_f(8)=t_{11}$, $w_f(9)=t_{32}$, $w_f(10)=t_{30}$, so

$$w(f)=t_{30}^3t_{31}^2t_{32}^2t_{10}t_{11}^2.$$

The (augmented) *cycle index* or *cycle indicator* $\tilde{Z}_n(t_{jk})$ of the *symmetric semigroup* $\Lambda_n=[n]^{[n]}$ of all functions $f:[n]\to[n]$ is the polynomial defined by

$$\tilde{Z}_n(t_{jk})=\sum_{f\in\Lambda_n}w(f).$$

For instance,

$$\tilde{Z}_2=t_{10}^2+t_{20}^2+2t_{10}t_{11}.$$

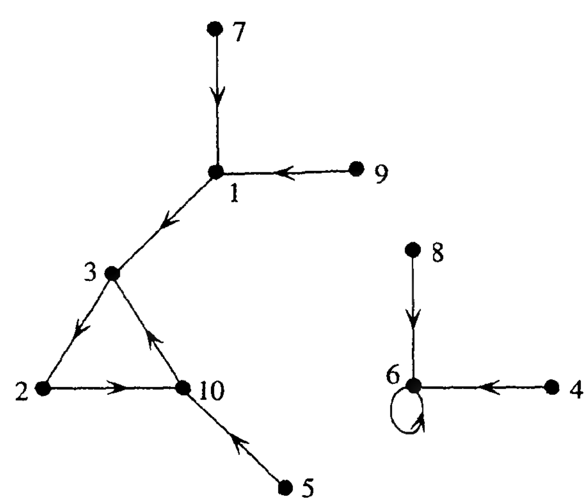


Figure 5-18. The digraph D_f of a function $f:[10]\to[10]$.

Note that

$$\tilde{Z}_n(t_{jk})|_{t_{jk}=0 \text{ for } k>0} = \tilde{Z}(\mathfrak{S}_n, t_{10}, t_{20}^2, t_{30}^3, \ldots),$$

where $\tilde{Z}(\mathfrak{S}_n)$ is defined in Example 5.2.10.

Show that

$$\sum_{n\geq 0} \tilde{Z}_n(t_{jk})\frac{x^n}{n!} = \exp\sum_{j\geq 1}\frac{1}{j}\left[t_{j0}xe^{t_{j1}xe^{t_{j2}xe^{\cdot^{\cdot^{\cdot}}}}}\right]^j. \tag{5.101}$$

b. [1+] Put each $t_{jk}=1$ to deduce (with $0^0=1$) that

$$\sum_{n\geq 0} n^n\frac{x^n}{n!} = \left[1 - xe^{xe^{xe^{\cdot^{\cdot^{\cdot}}}}}\right]^{-1}$$

$$= \left(1 - \sum_{n\geq 1} n^{n-1}\frac{x^n}{n!}\right)^{-1}.$$

c. [2] Fix $a, b \in \mathbb{P}$. Let $g(n)$ denote the number of functions $f : [n] \to [n]$ satisfying $f^a = f^{a+b}$ (exponents denote functional composition). Show that

$$\sum_{n\geq 0} g(n)\frac{x^n}{n!} = \exp\sum_{j|b}\frac{1}{j}\left[\underbrace{xe^{xe^{xe^{\cdot^{\cdot^{\cdot^{xe^x}}}}}}}_{a\ e\text{'s}}\right]^j \tag{5.102}$$

In particular, if $a=1$ then

$$\sum_{n\geq 0} g(n)\frac{x^n}{n!} = \exp\sum_{j|b}\frac{1}{j}(xe^x)^j. \tag{5.103}$$

d. [2] Deduce from (a) or (c) that the number $h(n)$ of functions $f : [n] \to [n]$ satisfying $f = f^{1+b}$ for *some* $b \in \mathbb{P}$ is given by

$$h(n) = \sum_{k=1}^{n} k^{n-k}(n)_k, \tag{5.104}$$

while the number $g(n)$ of idempotent functions $f : [n] \to [n]$ (i.e., $f^2 = f$) is given by

$$g(n) = \sum_{k=1}^{n} k^{n-k}\binom{n}{k}. \tag{5.105}$$

e. [2–] How many functions $f : [n] \to [n]$ satisfy $f^a = f^{a+1}$ for some $a \in \mathbb{P}$?

f. [1+] How many functions $f : [n] \to [n]$ have no fixed points?

5.33. [2] Find the flaw in the following argument. Let $c(n)$ be the total number of chains $\hat{0}=x_0 < x_1 < \cdots < x_k = \hat{1}$ in Π_n. Thus from Chapter 3.6,

$$c(n) = (2-\zeta)^{-1}(\hat{0}, \hat{1}),$$

where ζ is the zeta function of Π_n. Since

$$(2-\zeta)(x, y) = \begin{cases} 1, & x = y \\ -1, & x < y, \end{cases}$$

we have

$$E_{2-\zeta}(x) = x - \sum_{n\geq 2} \frac{x^n}{n!} = 1 + 2x - e^x.$$

Thus by Theorem 5.1.11, the generating function

$$y := E_c(x) = \sum_{n\geq 1} c(n)\frac{x^n}{n!}$$

satisfies

$$1 + 2y - e^y = x.$$

Equivalently,

$$y = (1 + 2x - e^x)^{\langle -1\rangle},$$

which is the same as (5.27).

5.34. **a.** [2] Fix $k \in \mathbb{P}$, and for $n \in \mathbb{N}$ define Ψ_n to be the subposet of Π_{kn+1} consisting of all partitions whose block sizes are $\equiv 1 \pmod{k}$. Thus Ψ_n is graded of rank n with rank function given by $\rho(\pi)=n - \frac{1}{k}(|\pi| - 1)$. Note that if $k=1$, then $\Psi_n = \Pi_{n+1}$. It is easy to see that if $\sigma \leq \pi$ in Ψ_n, then

$$[\sigma, \pi] \cong \Psi_0^{a_0} \times \Psi_1^{a_1} \times \cdots \times \Psi_n^{a_n}$$

for certain a_i satisfying $\sum i a_i = \rho(\sigma, \pi)$ (= the length of the interval $[\sigma, \pi]$) and $\sum a_i = |\pi|$. As in Section 5.1, we can define a *multiplicative function* $f : \mathbb{P} \to K$ on $\Psi = (\Psi_0, \Psi_1, \ldots)$, and the product (convolution) fg of two multiplicative functions. Lemma 5.1.10 remains true, so the multiplicative functions $f : \mathbb{P} \to K$ on Ψ form a monoid $M(\Psi) = M(\Psi, K)$.

As in Theorem 5.1.11, define a map $\varphi : M(\Psi) \to xK[[x]]$ by

$$\varphi(f) = \sum_{n\geq 0} f(n)\frac{x^{kn+1}}{(kn+1)!}.$$

Show that φ is an anti-isomorphism of monoids, so $\varphi(fg) = \varphi(g)(\varphi(f))$ (power series composition).

b. [1+] Let $q_n = \#\Psi_n$ and $\mu_n = \mu_{\Psi_n}(\hat{0}, \hat{1})$. Show that

$$\sum_{n\geq 0} q_n \frac{x^{kn+1}}{(kn+1)!} = e_k(e_k(x))$$

$$\sum_{n\geq 0} \mu_n \frac{x^{kn+1}}{(kn+1)!} = e_k^{\langle -1\rangle}(x),$$

where $e_k(x) = \sum_{n\geq 0} x^{kn+1}/(kn+1)!$. In particular, when $k=2$, $e_k(x) = \sinh x$.

c. [2] Let $\chi_n(t)$ denote the characteristic polynomial of Ψ_n (as defined in Section 3.10). Show that

$$\sum_{n\geq 0} \chi_n(t) \frac{x^{kn+1}}{(kn+1)!} = t^{-1/k} e_k\big(t^{1/k} e_k^{\langle -1\rangle}(x)\big). \tag{5.106}$$

Deduce that when $k=2$,

$$\chi_n(t) = (t-1^2)(t-3^2)\cdots[t-(2n-1)^2]. \tag{5.107}$$

In particular, $\mu_n = (-1)^n[1\cdot 3\cdot 5\cdots(2n-1)]^2$.

5.35. In this exercise we develop a noncrossing analogue of the exponential formula (Corollary 5.1.6) and its interpretation in terms of incidence algebras (Theorem 5.1.11).

a. [2+] Show that the number of noncrossing partitions of $[n]$ (as defined in Exercise 3.68) of type $s_1, \ldots, s_n$ (i.e., with s_i blocks of size i) is equal to $(n)_{k-1}/s_1!\cdots s_n!$, where $k = \sum s_i$.

b. [2+] Let NC_n denote the poset (actually a lattice) of noncrossing partitions of $[n]$, as defined in Exercise 3.68 (where $P_{1,n}$ is used instead of NC_n). Let K be a field. Given a function $f : \mathbb{P} \to K$, define a new function $h : \mathbb{P} \to K$ by

$$h(n) = \sum_{\pi=\{B_1,\ldots,B_k\}\in \mathrm{NC}_n} f(\#B_1)f(\#B_2)\cdots f(\#B_k).$$

Let $F(x) = 1 + \sum_{n\geq 1} f(n)x^n$ and $H(x) = 1 + \sum_{n\geq 1} h(n)x^n$. Show that

$$xH(x) = \left(\frac{x}{F(x)}\right)^{\langle -1\rangle}. \tag{5.108}$$

c. [3–] Let $\mathbf{NC} = (\mathrm{NC}_2, \mathrm{NC}_3, \ldots)$. For each $n \geq 2$, let $f_n \in I(\mathrm{NC}_n, K)$, the incidence algebra of NC_n. It is easy to see that every interval $[\sigma, \pi]$ of NC_n has a canonical decomposition

$$[\sigma,\pi] \cong \mathrm{NC}_2^{a_2} \times \mathrm{NC}_3^{a_3} \times \cdots \times \mathrm{NC}_n^{a_n}, \tag{5.109}$$

where $|\sigma| - |\pi| = \sum(i-1)a_i$. Suppose that the sequence $f = (f_2, f_3, \ldots)$ satisfies the following property: there is a function (also denoted f) $f : \mathbb{P} \to K$ such that if $\sigma \leq \pi$ in NC_n and $[\sigma, \pi]$ satisfies (5.109), then

$$f_n(\sigma,\pi) = f(2)^{a_2} f(3)^{a_3} \cdots f(n)^{a_n}.$$

We then call f a *multiplicative function* on $\mathbf{NC}$. (This definition is in exact analogy with the definition of a multiplicative function on $\mathbf{\Pi}$ following Corollary 5.1.9.)

Let $M(\mathbf{NC})$ denote the set of all multiplicative functions on $\mathbf{NC}$. Define the convolution fg of $f, g \in M(\mathbf{NC})$ analogously to (5.12). It is not hard to see that $fg \in M(\mathbf{NC})$. Given $f \in M(\mathbf{NC})$, set $f(1) = 1$ and define

$$\Gamma_f(x) = \frac{1}{x}\left(\sum_{n\geq 1} f(n)x^n\right)^{\langle -1\rangle}.$$

Show that $\Gamma_{fg} = \Gamma_f \Gamma_g$ for all $f, g \in M(\mathrm{NC})$. (In particular, $M(\mathrm{NC})$ is a *commutative* monoid. This fact also follows by reasoning as in Exercise 3.65 and using the fact that every interval of NC_n is self-dual.)

5.36. **a.** [2+] Find the coefficients of the power series

$$y = \left[\tfrac{1}{2}(1 + 2x - e^x)\right]^{\langle -1 \rangle} - \left[\log(1 + 2x) - x\right]^{\langle -1 \rangle}.$$

b. [1+] Let $t(n)$ be the number of total partitions of n, as defined in Example 5.2.5. Let $g(n)$ have the same meaning as in Exercise 5.26. Deduce from (a) that $g(n) = 2^n t(n)$ for $n \geq 1$.

c. [2+] Give a simple combinatorial proof of (b).

5.37. **a.** [2+] Let $1 = p_0(x), p_1(x), \ldots$ be a sequence of polynomials (with coefficients in some field K of characteristic 0), with $\deg p_n = n$ for all $n \in \mathbb{N}$. Show that the following four conditions are equivalent:

(i) $p_n(x + y) = \sum_{k \geq 0} \binom{n}{k} p_k(x) p_{n-k}(y)$, for all $n \in \mathbb{N}$.

(ii) There exists a power series $f(u) = a_1 u + a_2 u^2 + \cdots \in K[[u]]$ such that

$$\sum_{n \geq 0} p_n(x) \frac{u^n}{n!} = \exp x f(u). \tag{5.110}$$

NOTE: The hypothesis that $\deg p_n = n$ implies that $a_1 \neq 0$.

(iii) $\sum_{n \geq 0} p_n(x) \frac{u^n}{n!} = \left(\sum_{n \geq 0} p_n(1) \frac{u^n}{n!}\right)^x$.

(iv) There exists a linear operator Q on the vector space $K[x]$ of all polynomials in x, with the following properties:

- Qx is a nonzero constant
- Q is a *shift-invariant operator*, i.e., for all $a \in K$, Q commutes with the *shift operator* E^a defined by $E^a p(x) = p(x + a)$.
- We have

$$Q p_n(x) = n p_{n-1}(x) \qquad \text{for all } n \in \mathbb{P}. \tag{5.111}$$

NOTE: A sequence $p_0, p_1, \ldots$ of polynomials satisfying the above conditions is said to be of *binomial type*. The operator Q is called a *delta operator*, and the (unique) sequence $1 = p_0(x), p_1(x), \ldots$ satisfying (5.111) is called a *basic sequence* for Q.

b. [3–] Show that the following sequences are of binomial type (with $p_0(x) = 1$ and with $n \geq 1$ below):

$$p_n(x) = x^n$$

$$p_n(x) = (x)_n = x(x-1)\cdots(x-n+1)$$

$$p_n(x) = x^{(n)} = x(x+1)\cdots(x+n-1)$$

$$p_n(x) = x(x - an)^{n-1} \quad \text{for fixed } a \in K \qquad \text{(Abel polynomials)}$$

$$p_n(x) = \sum_{k=1}^{n} S(n,k) x^k \qquad \text{(exponential polynomials)}$$

$$p_n(x) = \sum_{k=1}^{n} \frac{n!}{k!} \binom{n + (a-1)k - 1}{n-k} x^k \qquad \text{for fixed } a \in K$$

(Laguerre polynomials at $-x$, for $a = 1$)

$$p_n(x) = \sum_{k=1}^{n} \binom{n}{k} k^{n-k} x^k.$$

In each case, find the power series $f(u)$ of (a)(ii) above. What is the operator Q of (a)(iv)?

c. [2+] Let T be a shift-invariant operator, and let Q be a delta operator with basic sequence $p_n(x)$. Show that

$$T = \sum_{n \geq 0} a_n \frac{Q^n}{n!},$$

where

$$a_n = [T p_n(x)]_{x=0}.$$

d. [2+] Let Q be a delta operator with basic polynomials $p_n(x)$. Show that there exists a unique power series $q(u) = b_1 u + \cdots$ ($b_1 \neq 0$) satisfying $q(D) = Q$, where D is the shift-invariant operator d/dx. Show also that the power series $f(u)$ of (5.110) is given by $f(u) = q^{\langle -1 \rangle}(u)$.

e. [2+] Suppose that $1 = p_0, p_1, \ldots$ is a sequence of polynomials of binomial type. Let

$$q_n(x) = \frac{x}{x + \alpha n} p_n(x + \alpha n), \qquad n \geq 0,$$

where α is a parameter. Show that the sequence $q_0, q_1, q_2, \ldots$ is also a sequence of polynomials of binomial type.

5.38. a. [2–] Let P be a binomial poset with factorial function $B(n)$, and let $Z_n(x)$ be the zeta polynomial of an n-interval of P. (See Sections 3.11 and 3.15 for definitions.) Show that $n!Z_n(x)/B(n)$, $n \geq 0$, is a sequence of polynomials of binomial type, as defined in the previous exercise.

b. [2–] Let $\mathbf{Q} = (Q_1, Q_2, \ldots)$ be an exponential structure with denominator sequence $(M(1), M(2), \ldots)$, and let $P_n(r, t)$ be the polynomial (in t) of equation (5.74). Set $M(0) = 1$. Show that for fixed $r \in \mathbb{Z}$ (or even r an indeterminate), the sequence of polynomials $P_n(r, x)/M(n)$, $n \geq 0$, is a sequence of polynomials of binomial type. Note the special cases $r = 1$ (equation (5.72)) and $r = 0$ (equation (5.75)).

5.39. [2+] Let $f(n)$ be the number of partial orderings of $[n]$ which are isomorphic to posets P that can be obtained from a one-element poset by successive iterations of the operations $+$ (disjoint union) and $\oplus$ (ordinal sum). Such posets are called *series–parallel* posets. For instance, all 19 partial orderings of [3] are counted by $f(3)$. Let

$$F(x) = \sum_{n \geq 1} f(n) \frac{x^n}{n!} = x + 3\frac{x^2}{2!} + 19\frac{x^3}{3!} + 195\frac{x^4}{4!}$$
$$+ 2791\frac{x^5}{5!} + 51303\frac{x^6}{6!} + \cdots.$$

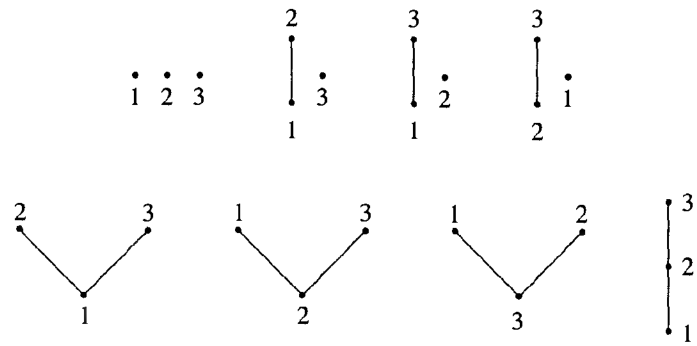


Figure 5-19. The eight inequivalent series–parallel posets on [3].

Show that

$$1 + F(x) = \exp\left[x + \frac{F(x)^2}{1 + F(x)}\right]. \tag{5.112}$$

Hence

$$F(x) = \left(\log(1+x) - \frac{x^2}{1+x}\right)^{\langle -1\rangle}$$
$$= \left(x - \tfrac{3}{2}x^2 + \tfrac{4}{3}x^3 - \tfrac{5}{4}x^4 + \tfrac{6}{5}x^5 - \cdots\right)^{\langle -1\rangle}.$$

5.40. a. [2+] Suppose that in the previous exercise we consider $P_1 \oplus P_2$ and $P_2 \oplus P_1$ to be equivalent. This induces an equivalence relation on the set of series-parallel posets on $[n]$. The equivalence classes are equivalent to what are called *series-parallel networks*. (The elements of the poset P correspond to the *edges* of a series–parallel network.) Figure 5-19 shows the eight inequivalent series-parallel posets on [3]. Let $s(n)$ be the number of equivalence classes of series-parallel posets on $[n]$ (or the number of series-parallel networks on n labeled edges), and set

$$S(x) = \sum_{n\geq 1} s(n)\frac{x^n}{n!}$$
$$= x + 2\frac{x^2}{2!} + 8\frac{x^3}{3!} + 52\frac{x^4}{4!} + 472\frac{x^5}{5!} + 5504\frac{x^6}{6!} + \cdots.$$

Show that

$$1 + S(x) = \exp\left(\tfrac{1}{2}[x + S(x)]\right). \tag{5.113}$$

Hence

$$S(x) = [2\log(1+x) - x]^{\langle -1\rangle}$$
$$= \left(x - x^2 + \tfrac{2}{3}x^3 - \tfrac{1}{2}x^4 + \tfrac{2}{5}x^5 - \tfrac{1}{3}x^6 + \cdots\right)^{\langle -1\rangle}.$$

b. [3–] Two graphs G_1 and G_2 (without loops or multiple edges) on the vertex set $[n]$ are said to be *switching-equivalent* if G_2 can be obtained from G_1 by choosing a subset X of $[n]$ and interchanging adjacency and non-adjacency between X and its complement $[n] - X$, leaving all edges within or outside X unchanged. Let $t(n)$ be the number of switching equivalence classes E of graphs on $[n]$ such that no graph in E contains an induced pentagon (5-cycle). Show that $t(n) = s(n-1)$.

c. [3–] A (real) *vector lattice* is a real vector space V with the additional structure of a lattice such that

$$\begin{aligned} x \le y \quad &\Longrightarrow \quad x + z \le y + z \quad \text{for all } x, y, z \in V \\ x \ge 0 \quad &\Longrightarrow \quad \alpha x \ge 0, \quad \text{for all } x \in V, \alpha \in \mathbb{R}^+. \end{aligned}$$

There is an obvious notion of isomorphism of vector lattices. Show that the number of non-isomorphic n-dimensional vector lattices is equal to the number of non-isomorphic *unlabeled* equivalence classes (as defined in (a)) of n-element series-parallel posets.

5.41. a. [2+] A tree on a linearly ordered vertex set is *alternating* (or *intransitive*) if for every vertex i the vertices adjacent to i are either all smaller than i or all larger than i. Let $f(n)$ denote the number of alternating trees on the vertex set $\{0, 1, \ldots, n\}$, and set

$$\begin{aligned} F(x) &= \sum_{n \ge 0} f(n) \frac{x^n}{n!} \\ &= 1 + x + 2\frac{x^2}{2!} + 7\frac{x^3}{3!} + 36\frac{x^4}{4!} + 246\frac{x^5}{5!} \\ &\quad + 2104\frac{x^6}{6!} + 21652\frac{x^7}{7!} + \cdots. \end{aligned}$$

Show that $F(x)$ satisfies the functional equation

$$F(x) = \exp\left(\frac{x}{2}[F(x) + 1]\right).$$

(Compare the similar but apparently unrelated (5.113).)

b. [2] Deduce that

$$f(n) = \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k} (k+1)^{n-1}.$$

c. [2] Let $f_k(n)$ denote the number of alternating trees on $\{0, 1, \ldots, n\}$ such that vertex 0 has degree k. Set

$$P_n(q) = \sum_{k=1}^{n} f_k(n) q^k.$$

For instance,

$$P_0(q) = 1, \qquad P_1(q) = q, \qquad P_2(q) = q^2 + q,$$
$$P_3(q) = q^3 + 3q^2 + 3q.$$

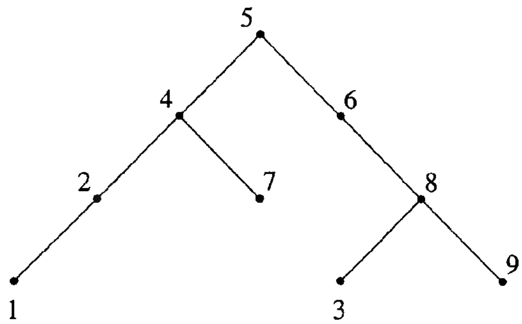


Figure 5-20. A local binary search tree.

Show that

$$\sum_{n\geq 0} P_n(q)\frac{x^n}{n!} = F(x)^q.$$

d. [2+] Show that

$$P_n(q) = \frac{q}{2^n}\sum_{k=0}^{n}\binom{n}{k}(q+k)^{n-1}.$$

e. [3] Show that if z is a complex number for which $P_n(z)=0$, then either $z=0$ or $\Re(z)=-n/2$, where $\Re$ denotes real part.

f. [2] Deduce from (e) that if $Q_n(q)=P_n(q)/q$, then

$$Q_n(q)=(-1)^{n-1}Q_n(-q-n).$$

g. [3–] A *local binary search tree* is a binary tree, say with vertex set $[n]$, such that every left child of a vertex is less than its parent, and every right child is greater than its parent. An example of such a tree is shown in Figure 5-20. Show that $f(n)$ is equal to the number of local binary search trees with vertex set $[n]$.

h. [3] Let $\mathcal{L}_n$ denote the set of all hyperplanes $x_i - x_j = 1$, $1 \leq i < j \leq n$, in $\mathbb{R}^n$. Show that the number of regions of $\mathcal{L}_n$ (i.e., the number of connected components of the space $\mathbb{R}^n - \bigcup_{H\in\mathcal{L}_n} H$) is equal to $f(n)$.

i. [3] Let L_n be the intersection poset of $\mathcal{L}_n$, as defined in Exercise 3.56. Show that the characteristic polynomial of L_n is given by

$$\chi(L_n, q) = (-1)^n P_n(-q).$$

This result generalizes (h), since by Exercise 3.56(a) the number of regions of $\mathcal{L}_n$ is equal to $|\chi(L_n, -1)|$.

j. [3–] An *alternating graph* on $[n]$ is a graph (without loops or multiple edges) on the vertex set $[n]$ such that every vertex is either smaller than all its neighbors or greater than all its neighbors. Let $g_k(n)$ denote the number of alternating graphs on $[n]$ with k edges. Show that

$$\sum_{n\geq 0}\sum_{k\geq 0} g_k(n)q^k\frac{x^n}{n!} = e^{-x}\sum_{n\geq 0}\left(\sum_{k=0}^{n}\binom{n}{k}_{q+1}\right)\frac{x^n}{n!},$$

where $\binom{n}{k}_{q+1}$ denotes the q-binomial coefficient $\binom{n}{k}$ with the variable q replaced by $q+1$.

k. [2+] An *edge-labeled alternating tree* is a tree, say with $n+1$ vertices, whose edges are labeled $1, 2, \ldots, n$ such that no path contains three consecutive edges whose labels are increasing. How many edge-labeled alternating trees have $n+1$ vertices?

5.42. **a.** [2] Let $y = R(x) = \sum_{n \geq 1} n^{n-1} \frac{x^n}{n!}$. Show from $y = xe^y$ that

$$[1 - R(x)]^{-1} = 1 + \sum_{n \geq 1} n^n \frac{x^n}{n!}.$$

b. [2+] Give a combinatorial proof, based on the fact that n^{n-1} is the number of rooted trees and n^n the number of double rooted trees on $[n]$.

5.43. [2] Generalize the bijection of Example 5.2.6 to show the following. Fix a sequence $(r_1, r_2, \ldots)$, with $r_i \in \mathbb{N}$ and $\sum i r_i = n < \infty$. Let $k = n + 1 - \sum r_i$. Then the number of (unordered) rooted trees with $n+1$ vertices and k leaves (or endpoints), whose leaves are labeled with the integers $1, 2, \ldots, k$, and with r_i nonleaf vertices of degree ($=$ number of successors) i, is equal to the number of partitions of the set $[n]$ into $n+1-k$ blocks, with r_i blocks of cardinality i.

5.44. [3–] Let $a_1, a_2, \ldots, a_k$ be positive integers summing to n. Let $f(a_1, \ldots, a_k)$ be the number of permutations $w_1 w_2 \cdots w_n$ of the multiset $\{1^{a_1}, \ldots, k^{a_k}\}$ such that if there is a subsequence of the form $xyyx$, then there must be an x between the two y's. More precisely, if $r < s < t < u$, $w_r = w_u$, and $w_s = w_t \neq w_r$, then there is a $s < v < t$ with $w_r = w_v$. Show that $f(a_1, \ldots, a_k) = n!/(n-k+1)!$.

5.45. [2+] A *recursively labeled tree* is a tree on the vertex set $[n]$, regarded as a poset with root $\hat{1}$, such that the vertices of every principal order ideal consist of consecutive integers. See Figure 5-21 for an example. Similarly define a *recursively labeled forest*. Let t_n (respectively, f_n) denote the number of recursively labeled trees (respectively, forests) on the vertex set $[n]$. Show that

$$t_n = \frac{1}{n}\binom{3n-2}{n-1}, \qquad f_n = \frac{1}{2n+1}\binom{3n}{n}.$$

Note that by Theorem 5.3.10 or Proposition 6.2.2, f_n is the number of plane ternary trees with $3n+1$ vertices (or, by removing the endpoints, the number of

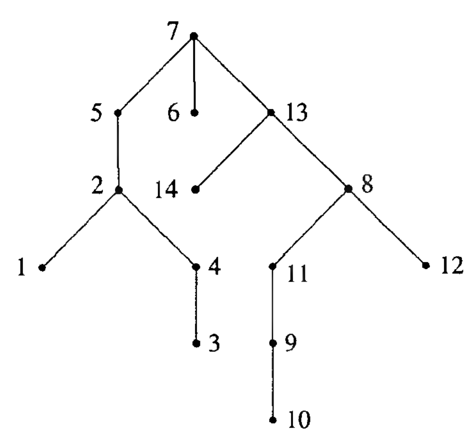


Figure 5-21. A recursively labeled tree.

ternary trees with n vertices). Similarly it is not hard to see that t_n is the number of ternary trees on n vertices except that the root has only two (linearly ordered) subtrees (rather than three). Equivalently, t_n is the number of ordered pairs of ternary trees with a total of $n-1$ vertices.

5.46. [2+] A tree on a linearly ordered vertex set is called *noncrossing* if ik and jl are not both edges whenever $i < j < k < l$. Show that the number $f(n)$ of noncrossing trees on $[n]$ is equal to $\frac{1}{2n-1}\binom{3(n-1)}{n-1}$, which by Theorem 5.3.10 or Proposition 6.2.2 is the number of ternary trees with $n-1$ vertices.

5.47. **a.** [2+] Show that the number of ways to write the cycle $(1, 2, \ldots, n) \in \mathfrak{S}_n$ as a product of $n-1$ transpositions (the minimum possible) is n^{n-2}. For instance (multiplying right to left), $(1,2,3) = (1,2)(2,3) = (2,3)(1,3) = (1,3)(1,2)$.

b. [3–] Define two factorizations of $(1, 2, \ldots, n)$ into $n-1$ transpositions to be *equivalent* if one can be obtained from the other by allowing transpositions with no common elements to commute. Thus the three factorizations of $(1,2,3)$ are all inequivalent, while the factorization $(1,5)(2,4)(2,3)(1,4)$ of $(1,2,3,4,5)$ is equivalent to itself and $(2,4)(1,5)(2,3)(1,4)$, $(1,5)(2,4)(1,4)(2,3)$, $(2,4)(1,5)(1,4)(2,3)$, and $(2,4)(2,3)(1,5)(1,4)$. Show that the number $g(n)$ of equivalence classes is equal to the number of noncrossing trees on the vertex set $[n]$, as defined in Exercise 5.46, and hence is equal to $\frac{1}{2n-1}\binom{3(n-1)}{n-1}$.

c. [3+] Let $\lambda = (\lambda_1, \lambda_2, \ldots)$ be a partition of n, and let w be a permutation of $1, 2, \ldots, n$ of cycle type λ. Let $f(\lambda)$ be the number of ways to write $w = t_1 t_2 \cdots t_k$ where the t_i's are transpositions that generate all of $\mathfrak{S}_n$, and where k is minimal with respect to the condition on the t_i's. (It is not hard to see that $k = n + \ell(\lambda) - 2$, where $\ell(\lambda)$ denotes the number of parts of λ.) Show that (writing ℓ for $\ell(\lambda)$)

$$f(\lambda) = (n + \ell - 2)!\, n^{\ell-3} \prod_{i=1}^{\ell} \frac{\lambda_i^{\lambda_i+1}}{\lambda_i!}.$$

5.48. **a.** [3–] Let τ be a rooted tree with vertex set $[n]$ and root 1. An *inversion* of τ is a pair (i, j) such that $1 < i < j$ and the unique path in τ from 1 to i passes through j. For instance, the tree τ of Figure 5-22 has the inversions $(3,4)$, $(2,4)$, $(2,6)$, and $(5,6)$. Let $\mathrm{inv}(\tau)$ denote the number of inversions

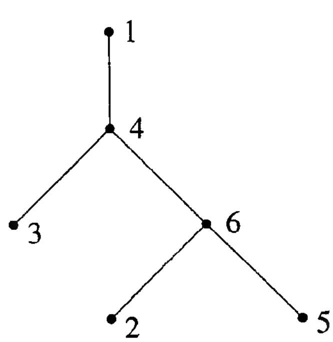


Figure 5-22. A tree with four inversions.

of τ. Define

$$I_n(t) = \sum_{\tau} t^{\mathrm{inv}(\tau)}, \tag{5.114}$$

summed over all n^{n-2} trees on $[n]$ with root 1. For instance,

$$\begin{aligned}
I_1(t) &= 1\\
I_2(t) &= 1\\
I_3(t) &= 2+t\\
I_4(t) &= 6+6t+3t^2+t^3\\
I_5(t) &= 24+36t+30t^2+20t^3+10t^4+4t^5+t^6\\
I_6(t) &= 120+240t+270t^2+240t^3+180t^4+120t^5+70t^6+35t^7\\
&\quad +15t^8+5t^9+t^{10}.
\end{aligned}$$

Show that

$$t^{n-1}I_n(1+t) = \sum_G t^{e(G)},$$

summed over all *connected* graphs G (without loops or multiple edges) on the vertex set $[n]$, where $e(G)$ is the number of edges of G.

It follows by a simple application of the exponential formula (Corollary 5.1.6) that

$$\sum_{n\geq 0}(1+t)^{\binom{n}{2}}\frac{x^n}{n!} = \exp \sum_{n\geq 1} t^{n-1}I_n(1+t)\frac{x^n}{n!}, \tag{5.115}$$

so

$$\sum_{n\geq 1} I_n(t)\frac{x^n}{n!} = (t-1)\log \sum_{n\geq 0} t^{\binom{n}{2}}(t-1)^{-n}\frac{x^n}{n!}.$$

b. [2] Deduce from (5.115) that

$$\sum_{n\geq 0} I_{n+1}(t)(t-1)^n\frac{x^n}{n!} = \frac{\sum_{n\geq 0} t^{\binom{n+1}{2}}\frac{x^n}{n!}}{\sum_{n\geq 0} t^{\binom{n}{2}}\frac{x^n}{n!}}.$$

5.49. a. [2] There are n parking spaces $1, 2, \ldots, n$ (in that order) on a one-way street. Cars $C_1, \ldots, C_n$ enter the street in that order and try to park. Each car C_i has a preferred space a_i. A car will drive to its preferred space and try to park there. If the space is already occupied, the car will park in the next available space. If the car must leave the street without parking, then the process fails. If $\alpha = (a_1, \ldots, a_n)$ is a sequence of preferences that allows every car to park, then we call α a *parking function*. Show that a sequence $(a_1, \ldots, a_n) \in [n]^n$ is a parking function if and only if the increasing rearrangement $b_1 \leq b_2 \leq \cdots \leq b_n$ of $a_1, a_2, \ldots, a_n$ satisfies $b_i \leq i$. In other words, $\alpha = (a_1, \ldots, a_n)$ is a parking function if and only if the sequence

$(a_1 - 1, \ldots, a_n - 1)$ is a permutation of the inversion table of a permutation $\pi \in \mathfrak{S}_n$, as defined in Section 1.3.

b. [2+] Regard the elements of the group $G = \mathbb{Z}/(n+1)\mathbb{Z}$ as being the integers $0, 1, \ldots, n$. Let H be the (cyclic) subgroup of order $n+1$ of the group G^n generated by $(1, 1, \ldots, 1)$. Show that each coset of H contains exactly one parking function. Hence the number $P(n)$ of parking functions of length n is given by

$$P(n) = (n+1)^{n-1}. \tag{5.116}$$

c. [3–] Let $\mathcal{P}_n$ denote the set of all parking functions $\alpha = (a_1, \ldots, a_n)$ of length n, and write $|\alpha| = a_1 + \cdots + a_n$. Show that

$$\sum_{\alpha \in \mathcal{P}_n} t^{|\alpha|} = t^{\binom{n+1}{2}} I_{n+1}(1/t),$$

where $I_n(t)$ is defined in equation (5.114). Try to give a bijective proof. (Note also that putting $t = 1$ yields (5.116).)

d. [2] Let $\alpha = (a_1, \ldots, a_n)$ be a parking function. Suppose that when the cars $C_1, \ldots, C_n$ park according to α, then C_i occupies space $w(i)$. Hence w is a permutation of $1, 2, \ldots, n$, which we denote by $w(\alpha)$. For instance, $w(3, 1, 3, 5, 1, 3) = 314526$. Given $u = u_1 \cdots u_n \in \mathfrak{S}_n$, let $\nu(u)$ be the number of parking functions α for which $w(\alpha) = u$. For $1 \leq j \leq n$, define

$$\tau(u, j) = 1 + \max\{k : j-1, j-2, \ldots, j-k \text{ precede } j \text{ in } u\},$$

and set $\tau(u) = (\tau(u, 1), \ldots, \tau(u, n))$. For instance, $\tau(314526) = (1, 2, 1, 2, 3, 6)$. Show that

$$\nu(u) = \tau(u, 1) \cdots \tau(u, n).$$

e. [3–] Given $\sigma \in \mathbb{P}^n$, let

$$T_\sigma = \{u \in \mathfrak{S}_n : \tau(u) = \sigma\}.$$

For instance, $T_{(1,2,1,2,1)} = \{53412, 35412, 53142, 35142, 31542, 51342, 15342, 13542\}$. Suppose that $\sigma = (s_1, \ldots, s_n) = \tau(u)$ for some $u \in \mathfrak{S}_n$. (For the characterization and enumeration of the sequences $\tau(u)$, $u \in \mathfrak{S}_n$, see Exercise 6.19(z).) Define

$$t_i = \max\{j : s_{i+r} \leq r \text{ for } 1 \leq r \leq j\}.$$

(If $s_{i+1} > 1$ then set $t_i = 0$.) Show that

$$\#T_\sigma = \frac{n!}{(s_1 + t_1)(s_2 + t_2) \cdots (s_n + t_n)}.$$

f. [3–] A parking function $\alpha = (a_1, \ldots, a_n)$ is said to be *prime* if for all $1 \leq j \leq n-1$, at least $j+1$ cars want to park in the first j places. (Equivalently, if we remove some term of α equal to 1, then we still have a parking function.) Show that the number $Q(n)$ of prime parking functions of length n is equal to $(n-1)^{n-1}$.

5.50. **a.** [3–] Let $\mathcal{S}_n$ denote the set of all hyperplanes $x_i - x_j = 0, 1$ $(1 \le i < j \le n)$ in $\mathbb{R}^n$. (Hence $\#\mathcal{S}_n = n(n-1)$.) Show bijectively that the number of regions of $\mathcal{S}_n$ (i.e., the number of connected components of the space $\mathbb{R}^n - \bigcup_{H \in \mathcal{S}_n} H$) is equal to $(n+1)^{n-1}$.

b. [2+] Let $\mathcal{A}$ be a finite hyperplane arrangement in $\mathbb{R}^n$ with intersection poset L, as in Exercise 3.56. Suppose that $\mathcal{A}$ is defined over $\mathbb{Z}$, i.e., the equations of the hyperplanes in $\mathcal{A}$ can be written with integer coefficients. If p is a prime, then we define $\mathcal{A}_p$ to be the arrangement $\mathcal{A}$ "reduced modulo p," i.e., regard the equations of the hyperplanes in $\mathcal{A}$ as being defined over the field $\mathbb{F}_p$. Hence $\mathcal{A}_p$ is an arrangement of hyperplanes in $\mathbb{F}_p^n$. Let $\chi(L, q)$ denote the characteristic polynomial of L. Show that for p sufficiently large, we have

$$\chi(L, p) = \#\left(\mathbb{F}_p^n - \bigcup_{H \in \mathcal{A}_p} H\right).$$

c. [3–] Let $L_{\mathcal{S}_n}$ denote the intersection poset of $\mathcal{S}_n$. Use (b) to show that

$$\chi(L_{\mathcal{S}_n}, q) = q(q-n)^{n-1}.$$

This result generalizes (a), since by Exercise 3.56(a) the number of regions of $\mathcal{S}_n$ is equal to $|\chi(L_{\mathcal{S}_n}, -1)|$.

d. [3] Let R_0 be the region of $\mathcal{S}_n$ defined by $x_i - 1 < x_j < x_i$ for all $i < j$. For any region R of $\mathcal{S}_n$, let $d(R)$ be the number of hyperplanes $H \in \mathcal{S}_n$ that *separate* R from R_0, i.e., R and R_0 lie on different sides of H. Define the polynomial

$$J_n(q) = \sum_R q^{d(R)},$$

summed over all regions of $\mathcal{S}_n$. Show that

$$J_n(q) = q^{\binom{n}{2}} I_{n+1}(1/q),$$

where $I_n(t)$ is defined in equation (5.114).

e. [2+] Show that (d) is equivalent to the following result. Given a permutation $\pi \in \mathfrak{S}_n$, let $P_\pi = \{(i,j) : 1 \le i < j \le n,\ \pi(i) < \pi(j)\}$. Define a partial ordering on P_π by $(i,j) \le (k,l)$ if $k \le i < j \le l$. Let $F(J(P_\pi), q)$ denote the rank-generating function of the lattice of order ideals of P_π. (For instance, if $\pi = n, n-1, \ldots, 1$, then $P_\pi = \emptyset$ and $F(J(P_\pi), q) = 1$.) Then

$$\sum_{\pi \in \mathfrak{S}_n} F(J(P_\pi), q) = I_{n+1}(q).$$

f. [3–] Show that the number of elements of rank k in the intersection poset $L_{\mathcal{S}_n}$ is equal to the number of ways to partition the set $[n]$ into $n-k$ blocks, and

then linearly order each block. (It is easy to see that this number is given by

$$\frac{n!}{(n-k)!}\binom{n-1}{k};$$

see Exercise 5.17.)

5.51. [2+] Let $A(x)=ax+\cdots$, $B(x)=bx+\cdots$, $C(x)=c+\cdots \in K[[x]]$ with $abc \neq 0$. Show that the following two formulas are equivalent:

$$\text{(i)}\ \ A(x)^{\langle -1\rangle} = C(x)B(x)^{\langle -1\rangle}$$

$$\text{(ii)}\ \ \frac{x}{C(A(x))} = [xC(B(x))]^{\langle -1\rangle}.$$

5.52. a. [2] Let $F(x)=x+\sum_{n\geq 2} f_n \frac{x^n}{n!} \in K[[x]]$. Given $k \in \mathbb{P}$, let

$$F^{\langle k\rangle}(x) = x + \sum_{n\geq 2} \varphi_n(k)\frac{x^n}{n!}. \tag{5.117}$$

Show that for fixed n, the function $\varphi_n(k)$ is a polynomial in k (whose coefficients are polynomials in $f_2, \ldots, f_n$). For instance,

$$\begin{aligned}
\varphi_2(k) &= f_2 k \\
\varphi_3(k) &= f_3 k + 3f_2^2\binom{k}{2} \\
\varphi_4(k) &= f_4 k + \left(10 f_2 f_3 + 3 f_2^3\right)\binom{k}{2} + 18 f_2^3\binom{k}{3} \\
\varphi_5(k) &= f_5 k + \left(15 f_2 f_4 + 10 f_3^2 + 25 f_2^2 f_3\right)\binom{k}{2} \\
&\quad + \left(130 f_2^2 f_3 + 75 f_2^4\right)\binom{k}{3} + 180 f_2^4\binom{k}{4}.
\end{aligned}$$

b. [2] Since $\varphi_n(k)$ is a polynomial in k, it can be defined for any $k \in K$ (or for k an indeterminate). Thus (5.117) allows us to define $F^{\langle k\rangle}(x)$ for any k. Show that for all $j, k \in K$, we have

$$F^{\langle j+k\rangle}(x) = F^{\langle j\rangle}\left(F^{\langle k\rangle}(x)\right)$$

$$F^{\langle jk\rangle}(x) = \left(F^{\langle j\rangle}\right)^{\langle k\rangle}(x).$$

In particular, the two ways of defining $F^{\langle -1\rangle}(x)$ (viz., by putting $k=-1$ in (5.117), or as the compositional inverse of $F(x)$) agree.

c. [5–] Investigate the combinatorial significance of "fractional composition." For instance, setting

$$
\begin{aligned}
(e^x - 1)^{\langle 1/2 \rangle} &= \sum_{n\geq 1} a_n \frac{x^n}{n!} \\
&= x + \frac{1}{2}\frac{x^2}{2!} + \frac{1}{2^3}\frac{x^3}{3!} + \frac{1}{2^5}\frac{x^5}{5!} - \frac{7}{2^7}\frac{x^6}{6!} + \frac{1}{2^7}\frac{x^7}{7!} \\
&\quad + \frac{159}{2^8}\frac{x^8}{8!} - \frac{843}{2^8}\frac{x^9}{9!} - \frac{1231}{2^{12}}\frac{x^{10}}{10!} + \frac{2359233}{2^{14}}\frac{x^{11}}{11!} \\
&\quad - \frac{13303471}{2^{14}}\frac{x^{12}}{12!} - \frac{271566005}{2^{15}}\frac{x^{13}}{13!} \\
&\quad + \frac{10142361989}{2^{16}}\frac{x^{14}}{14!} + \frac{126956968965}{2^{18}}\frac{x^{15}}{15!} \\
&\quad - \frac{10502027401553}{2^{18}}\frac{x^{16}}{16!} + \cdots,
\end{aligned}
$$

do the coefficients a_n have a simple combinatorial interpretation? (Unfortunately, they are not integers, nor do their signs seem predictable.)

5.53. [2+] Find the sum of the first n terms in the binomial expansion of

$$\left(1 - \frac{1}{2}\right)^{-n} = 1 + \frac{1}{2}n + \frac{1}{4}\binom{n+1}{2} + \cdots.$$

For instance, when $n = 3$ we get $1 + \frac{3}{2} + \frac{6}{4} = 4$. (Use the Lagrange inversion formula.)

5.54. [2+] For each of the following four power series $F(x)$, find for all $n \in \mathbb{P}$ the coefficient of $1/x$ in the Laurent expansion about 0 of $F(x)^{-n}$: $\sin x$, $\tan x$, $\log(1+x)$, $1 + x - \sqrt{1+x^2}$.

5.55. **a.** [2] Find the unique power series $F_1(x) \in \mathbb{Q}[[x]]$ such that for all $n \in \mathbb{N}$, we have $[x^n]F_1(x)^{n+1} = 1$.

b. [2+] Find the unique power series $F_2(x) \in \mathbb{Q}[[x]]$ such that for all $n \in \mathbb{N}$, we have $[x^n]F_2(x)^{2n+1} = 1$.

c. [2+] Let $k \in \mathbb{P}$, $k \geq 3$. What difficulty arises in trying to find an explicit expression for the unique power series $F_k(x) \in \mathbb{Q}[[x]]$ such that for all $n \in \mathbb{N}$, we have $[x^n]F_k(x)^{kn+1} = 1$?

5.56. **a.** [2+] Let $F(x) = a_1 x + a_2 x^2 + \cdots \in K[[x]]$ with $a_1 \neq 0$, and let $n \in \mathbb{P}$. Show that

$$n[x^n] \log \frac{F^{\langle -1 \rangle}(x)}{x} = [x^n]\left(\frac{x}{F(x)}\right)^n. \tag{5.118}$$

(This formula may be regarded as the "correct" case $k = 0$ of (5.53).)

b. [2] Find the unique power series $G(x) = 1 + x - \frac{1}{2}x^2 + \cdots$ satisfying $G(0) = 1$, $[x]G(x) = 1$, and $[x^n]G(x)^n = 0$ for $n > 1$.

5.57. [2] Show that the coefficient of x^{n-1} in the power series expansion of the rational function $(1+x)^{2n-1}(2+x)^{-n}$ is equal to $\frac{1}{2}$. Equivalently, the unique power

series $J(x) \in \mathbb{Q}[[x]]$ satisfying

$$[x^{n-1}]\frac{J(x)^n}{1+x} = \frac{1}{2} \qquad \text{for all } n \in \mathbb{P}$$

is given by $J(x) = (1+x)^2/(2+x)$.

5.58. [3–] Let $f(x)$ and $g(x)$ be power series with $g(0) = 1$. Suppose that

$$f(x) = g(xf(x)^\alpha), \tag{5.119}$$

where α is a parameter (variable). Show that

$$(t + \alpha n)[x^n]f(x)^t = t[x^n]g(x)^{t+\alpha n},$$

as a polynomial identity in the two variables t and α.

5.59. [2+] Let $f(x) \in K[[x]]$ with $f(0) = 0$. Let $F(x, y) \in K[[x, y]]$, and suppose that f satisfies the functional equation $f = F(x, f)$. Show that for every $k \in \mathbb{P}$,

$$f(x)^k = \sum_{n \geq 1} \frac{k}{n}[y^{n-k}]F(x, y)^n.$$

5.60. a. [2] Let $A(x) = 1 + \sum_{n\geq 1} a_n x^n \in K[[x]]$. For fixed $k \in \mathbb{N}$, define for $n \in \mathbb{Z}$

$$q_k(n) = [x^k]A(x)^n.$$

Show that $q_k(n)$ is a polynomial in n of degree $\leq k$.

b. [2] Let $F(x) = x + \sum_{n\geq 2} f_n(x^n/n!) \in K[[x]]$ (where char $K = 0$). Define functions $p_k(n)$ by

$$e^{tF(x)} = \sum_{n\geq 0}\sum_{k\geq 0} p_k(n)t^n \frac{x^{n+k}}{(n+k)!}.$$

Show that $p_k(n)$ is a polynomial in n (of degree $\leq 2k$).

c. [2+] Let $F(x)$ and $p_k(n)$ be as in (b). Since $p_k(n)$ is a polynomial in n, it is defined for all $n \in \mathbb{Z}$. Show that

$$e^{tF^{\langle -1\rangle}(x)} = \sum_{n\geq 0}\sum_{k\geq 0}(-1)^k p_k(-n-k)t^n \frac{x^{n+k}}{(n+k)!}.$$

d. [2] What are $p_k(n)$ and $p_k(-n-k)$ in the special case $F(x) = e^x - 1$?

e. [2] Find $p_k(n)$ when $F(x) = x/(1-x)$. What does (c) tell us about $p_k(n)$?

f. [2] Find $p_k(n)$ when $F(x) = xe^{-x}$. Deduce a formula for

$$\exp t(xe^{-x})^{\langle -1\rangle} = \exp t\sum_{n\geq 1} n^{n-1}\frac{x^n}{n!}.$$

5.61. a. [2] Let P and Q be finite posets with $\hat{1}$'s. For any poset T let $\bar{T} = T \cup \{\hat{0}\}$, and for any finite poset T with $\hat{0}$ and $\hat{1}$ let $\mu(T) = \mu_T(\hat{0}, \hat{1})$. Show that

$$-\mu(\overline{P \times Q}) = \mu(\bar{P} \times \bar{Q}) = \mu(\bar{P})\mu(\bar{Q}).$$

b. [2] Use (a) and Corollary 5.5.5 to give a direct proof of equation (5.78).

5.62. a. [2+] Let $f_r(n)$ be the number of $n \times n$ $\mathbb{N}$-matrices A with every row and column sum equal to r and with at most two nonzero entries in every row

(and hence in every column [why?]). Find

$$\sum_{n\geq 0} f_r(n)\frac{x^n}{n!^2}.$$

b. [1] Use (a) to find $f_3(n)$ explicitly.

5.63. [2+] Let $N_k(n)$ denote the number of sequences $(P_1, P_2, \ldots, P_{2k})$ of $n \times n$ permutation matrices P_i such that each entry of $P_1 + P_2 + \cdots + P_{2k}$ is 0, k, or $2k$. Show that

$$\sum_{n\geq 0} N_k(n)\frac{x^n}{n!^2} = (1-x)^{-\binom{2k-1}{k}}\exp\left(x\left[1-\binom{2k-1}{k}\right]\right).$$

5.64. **a.** [2+] Let $\mathcal{D}_n$ be the set of all $n \times n$ matrices of $+1$'s and -1's. For $k \in \mathbb{P}$ let

$$f_k(n) = 2^{-n^2}\sum_{M\in\mathcal{D}_n}(\det M)^k$$

$$g_k(n) = 2^{-n^2}\sum_{M\in\mathcal{D}_n}(\operatorname{per} M)^k,$$

where per denotes the permanent function, defined by

$$\operatorname{per}(m_{ij}) = \sum_{\pi\in\mathfrak{S}_n} m_{1,\pi(1)}m_{2,\pi(2)}\cdots m_{n,\pi(n)}.$$

Find $f_k(n)$ and $g_k(n)$ explicitly when k is odd or $k = 2$.

b. [3–] Show that $f_4(n) = g_4(n)$, and show that

$$\sum_{n\geq 0} f_4(n)\frac{x^n}{n!^2} = (1-x)^{-3}e^{-2x}. \tag{5.120}$$

Hint. We have

$$\sum_M (\det M)^4 = \sum_M\left(\sum_{\pi\in\mathfrak{S}_n} \pm m_{1,\pi(1)}\cdots m_{n,\pi(n)}\right)^4.$$

Interchange the order of summation and use Exercise 5.63.

c. [2+] Show that $f_{2k}(n) < g_{2k}(n)$ if $k \geq 3$ and $n \geq 3$.

d. [3–] Let $\mathcal{D}'_n$ be the set of all $n \times n$ 0–1 matrices. Let $f'_k(n)$ and $g'_k(n)$ be defined analogously to $f_k(n)$ and $g_k(n)$. Show that $f'_k(n) = 2^{-kn}f_k(n+1)$. Show also that

$$g'_1(n) = 2^{-n}n!$$

$$g'_2(n) = 4^n n!^2\left(1+\frac{1}{1!}+\frac{1}{2!}+\cdots+\frac{1}{n!}\right).$$

5.65. **a.** [3–] Let $f(m, n)$ be the number of $m \times n$ $\mathbb{N}$-matrices with every row and column sum at most two. For instance, $f(1, n) = 1 + 2n + \binom{n}{2}$. Show that

$$\begin{aligned} F(x, y) &:= \sum_{m,n\geq 0} f(m,n)\frac{x^m y^n}{m!\,n!} \\ &= (1-xy)^{-\frac{1}{2}}\exp\left[\frac{\frac{1}{2}xy(3-xy)+\frac{1}{2}(x+y)(2-xy)}{1-xy}\right]. \end{aligned} \tag{5.121}$$

b. [2] Deduce from (a) that

$$\sum_{n\geq 0} f(n,n)\frac{t^n}{n!^2} = (1-t)^{-\frac{1}{2}}e^{\frac{t(3-t)}{2(1-t)}}\sum_{n\geq 0}\frac{t^n}{n!^2}\left(\frac{1-\frac{1}{2}t}{1-t}\right)^{2n}.$$

The latter sum may be rewritten as $J_0[(2-t)/\sqrt{t-1}]$, where J_0 denotes the Bessel function of order zero.

5.66. [2+] Let $\mathbf{L}=\mathbf{L}(K_{rs})$ be the Laplacian matrix of the complete bipartite graph K_{rs}.

a. Find a simple upper bound on rank$(\mathbf{L}-r\mathbf{I})$. Deduce a lower bound on the number of eigenvalues of $\mathbf{L}$ equal to r.

b. Assume $r\neq s$, and do the same as (a) for s instead of r.

c. Find the remaining eigenvalues of $\mathbf{L}$.

d. Use (a)–(c) to compute $c(K_{rs})$, the number of spanning trees of K_{rs}.

5.67. [3] Let q be a prime power, and let $\mathbb{F}_q$ denote the finite field with q elements. Given $f:\binom{[n]}{2}\to\mathbb{F}_q$ and a free tree T on the vertex set $[n]$, define $f(T)=\prod_e f(e)$, where e ranges over all edges (regarded as two-element subsets of $[n]$) of T. Let $P_n(q)$ denote the number of maps f for which

$$\sum_T f(T)\neq 0 \quad (\text{in } \mathbb{F}_q),$$

where T ranges over all n^{n-2} free trees on the vertex set $[n]$. Show that

$$P_{2m}(q) = q^{m(m-1)}(q-1)(q^3-1)\cdots(q^{2m-1}-1)$$
$$P_{2m+1}(q) = q^{m(m+1)}(q-1)(q^3-1)\cdots(q^{2m-1}-1).$$

5.68. [3–] This exercise assumes a basic knowledge of the character theory of finite abelian groups. Let Γ be a finite abelian group, written additively. Let $\hat{\Gamma}$ denote the set of (irreducible) characters $\chi:\Gamma\to\mathbb{C}$ of Γ, with the trivial character denoted by χ_0. Let $\sigma:\Gamma\to K$ be a weight function (where K is a field of characteristic zero). Define $D=D_\sigma$ to be the digraph on the vertex set Γ with an edge $u\to u+v$ of weight $\sigma(v)$ for all $u,v\in\Gamma$. Note that D is balanced as a weighted digraph (every vertex has indegree and outdegree equal to $\sum_{u\in\Gamma}\sigma(u)$). If T is any spanning subgraph of D, then let $\sigma(T)=\prod_e\sigma(e)$, where e ranges over all edges of T. Define

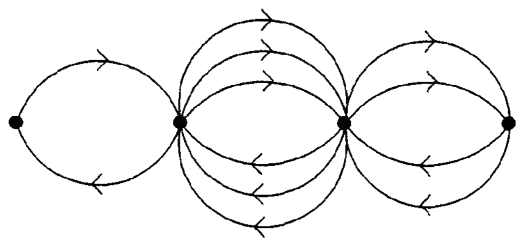
$$c_\sigma(D)=\sum_T\sigma(T),$$

where T ranges over all oriented (spanning) subtrees of D with a fixed root. Show that

$$c_\sigma(D)=\frac{1}{|\Gamma|}\prod_{\substack{\chi\in\hat{\Gamma}\\ \chi\neq\chi_0}}\sum_{v\in\Gamma}\sigma(v)[1-\chi(v)].$$

5.69. Choose positive integers $a_1,\ldots,a_{p-1}$. Let $D=D(a_1,\ldots,a_{p-1})$ be the digraph defined as follows. The vertices of D are $v_1,\ldots,v_p$. For each $1\leq i\leq p-1$,

there are a_i edges from x_i to x_{i+1} and a_i edges from x_{i+1} to x_i. For instance, $D(1, 3, 2)$ looks like



a. [2–] Find by a direct argument (no determinants) the number $\tau(D, v)$ of oriented subtrees with a given root v.

b. [2–] Find the number $\epsilon(D, e)$ of Eulerian tours of D whose first edge is e.

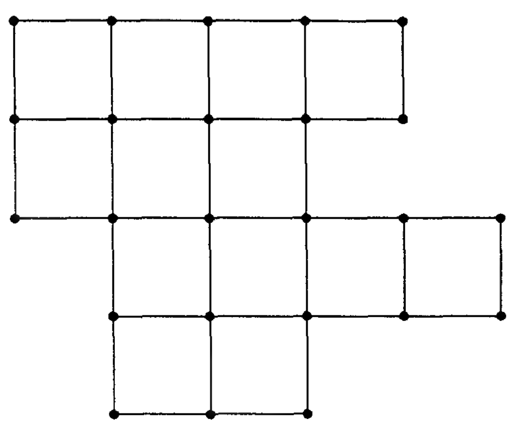
5.70. [2] Let $d > 1$. A *d-ary de Bruijn sequence* of degree n is a sequence $A = a_1 a_2 \cdots a_{d^n}$ whose entries a_i belong to $\{0, 1, \ldots, d-1\}$ such that every d-ary sequence $b_1 b_2 \cdots b_n$ of length n occurs exactly once as a circular factor of A. Find the number of d-ary de Bruijn sequences of degree n that begin with n 0's.

5.71. [2+] Let G be a regular graph of degree d with no loops or multiple edges. Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$ be nonzero real numbers such that for all $\ell \geq 1$, the number $W(\ell)$ of closed walks in G of length ℓ is given by

$$W(\ell) = \sum_{j=1}^{m} \lambda_j^{\ell}. \tag{5.122}$$

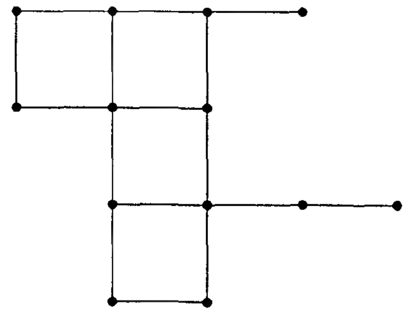
Find the number $c(G)$ of spanning trees of G in terms of the given data.

5.72. [3–] Let V be the subset of $\mathbb{Z} \times \mathbb{Z}$ on or inside some simple closed polygonal curve whose vertices belong to $\mathbb{Z} \times \mathbb{Z}$, such that every line segment that makes up the curve is parallel to either the x-axis or the y-axis. Draw an edge e between any two points of V at distance one apart, provided e lies on or inside the boundary curve. We obtain a planar graph G, an example being



Let G' be the dual graph G^* with the "outside" vertex deleted. (The vertices of G^* are the regions of G. For every edge e of G there is an edge e^* of G^* connecting the two regions that have e on their boundary.) For the above example,

G' is given by



Let $\lambda_1, \ldots, \lambda_p$ denote the eigenvalues of G' (i.e., of the adjacency matrix $\mathbf{A}(G')$). Show that

$$c(G) = \prod_{i=1}^{p}(4 - \lambda_i).$$

5.73. [5–] Let $\mathcal{B}(n)$ be the set of (binary) de Bruijn sequences of degree n, and let $\mathcal{S}_n$ be the set of all binary sequences of length 2^n. According to Corollary 5.6.15 we have $[\#\mathcal{B}(n)]^2 = \#\mathcal{S}(n)$. Find an explicit bijection $\mathcal{B}(n) \times \mathcal{B}(n) \to \mathcal{S}(n)$.

5.74. Let D be a digraph with p vertices, and let ℓ be a fixed positive integer. Suppose that for every pair u, v of vertices of D, there is a unique (directed) walk of length ℓ from u to v.

a. [2+] What are the eigenvalues of the (directed) adjacency matrix $\mathbf{A} = \mathbf{A}(D)$?

b. [2] How many loops (v, v) does D have?

c. [3–] Show that D is connected and balanced.

d. [1] Let d be the indegree and outdegree of each vertex of D. (By (c), all vertices have the same indegree and outdegree.) Find a simple formula relating p, d, and ℓ.

e. [2] How many Eulerian tours does D have starting with a given edge e?

f. [5–] What more can be said about D? Must D be a de Bruijn graph (the graphs used to solve Exercise 5.70)?

Solutions to Exercises

5.1. a. Let $h(n)$ be the desired number. By Proposition 5.1.3, we have

$$\begin{aligned}
E_h(x) &= \left(\sum_{n\geq 0}\frac{x^n}{n!}\right)^2 \left(\sum_{n\geq 0}\frac{x^{2n+1}}{(2n+1)!}\right)\left(\sum_{n\geq 0}\frac{x^{2n}}{(2n)!}\right)\\
&= e^{2x}\left(\frac{e^x - e^{-x}}{2}\right)\left(\frac{e^x + e^{-x}}{2}\right)\\
&= \frac{1}{4}(e^{4x} - 1)\\
&= \sum_{n\geq 1} 4^{n-1}\frac{x^n}{n!},
\end{aligned}$$

whence $h(n) = 4^{n-1}$.

b. Pick a set S of $2k$ poles to be either orange or purple, and pick a subset of S to be orange in 2^{2k} ways. Thus we obtain an extra factor of

$$\sum_{n\geq 0} 2^{2n}\frac{x^{2n}}{(2n)!} = \frac{1}{2}(e^{2x}+e^{-2x}).$$

Hence

$$E_h(x) = \tfrac{1}{4}(e^{4x}-1)\cdot\tfrac{1}{2}(e^{2x}+e^{-2x})$$

$$= \tfrac{1}{8}(e^{6x}-e^{-2x}),$$

so $h(n)=\frac{1}{8}[6^n-(-2)^n]$.

5.2. a. By Exercise 1.40, there are unique numbers a_i such that

$$1+\sum_{n\geq 1} f_n x^n = \prod_{i\geq 1}(1-x^i)^{-a_i}.$$

It is easily seen that $f_n\in\mathbb{Z}$ for all $n\in[N]$ if and only if $a_i\in\mathbb{Z}$ for all $i\in[N]$. Now by the solution to Exercise 1.40

$$h_n = \sum_{d|n} d a_d,$$

so by the classical Möbius inversion formula,

$$a_n = \frac{1}{n}\sum_{d|n} h_d\,\mu(n/d),$$

and the equivalence of (i) and (ii) follows.

Now let $p\mid n$, and let S be the set of distinct primes other than p which divide n. If $T\subseteq S$ then write $\Pi(T)=\prod_{q\in T} q$. Then

$$A_n := \sum_{d|n} h_d\,\mu(n/d) = \sum_{T\subseteq S}(-1)^{\#T}\left(h_{n/\Pi(T)} - h_{n/p\Pi(T)}\right). \qquad (5.123)$$

Hence if (iii) holds for *all* $n\in[N]$, then by (5.123) we have $p^r\mid A_n$. Thus $A_n\equiv 0 \pmod n$. Conversely, if (ii) holds for all $n\in[N]$ then (iii) follows from (5.123) by an easy induction on n.

Finally observe that

$$\exp\sum_{n\geq 1}\left(\sum_{i=1}^{N}\alpha_i^n\right)\frac{x^n}{n} = \frac{1}{(1-\alpha_1 x)\cdots(1-\alpha_N x)}.$$

From this it is immediate that (iv) $\Rightarrow$ (i). Conversely, if (i) holds then let

$$\left(1+\sum_{n\geq 1} f_n x^n\right)^{-1} = 1+\sum_{n\geq 1} e_n x^n.$$

Clearly $e_n\in\mathbb{Z}$ for $n\in[N]$. Set

$$1+\sum_{n=1}^{N} e_n t^n = \prod_{1}^{N}(1-\alpha_i t).$$

Then $h_n=\sum_{i=1}^N \alpha_i^n$ for all $n\in[N]$, as desired.

The equivalence of (ii) and (iv) goes back to W. Jänischen, *Sitz. Berliner Math. Gesellschaft* **20** (1921), 23–29. The condition (iii) is due to I. Schur, *Comp. Math.* 4 (1937), 432–444, who obtains several related results. The equivalence of (i) and (ii) in the case $N \to \infty$ appears in L. Carlitz, *Proc. Amer. Math. Soc.* **9** (1958), 32–33. Additional references are J. S. Frame, *Canadian J. Math.* **1** (1949), 303–304; G. Almkvist, The integrity of ghosts, preprint; A. Dold, *Inv. Math.* **74** (1983), 419–435.

b. Let us say that a solution $\alpha = (\alpha_1, \dots, \alpha_k)$ has *degree n* if n is the smallest integer for which $\alpha \in \mathbb{F}_{q^n}^k$. By a simple Möbius inversion argument, the number M_n of solutions of degree n is given by $M_n = \sum_{d|n} N_d\, \mu(n/d)$. Write $\alpha^j = (\alpha_1^j, \dots, \alpha_k^j)$. If α is a solution of degree n, then the k-tuples $\alpha, \alpha^q, \alpha^{q^2}, \dots, \alpha^{q^{n-1}}$ are distinct solutions of degree n. Hence M_n is divisible by n. Now use the equivalence of (i) and (ii) in (a). See for instance K. Ireland and M. Rosen, *A Classical Introduction to Modern Number Theory*, second ed., Springer-Verlag, New York/Berlin/Heidelberg, 1990 (§11.1).

A considerably deeper result, first proved by B. Dwork, *Amer. J. Math.* **82** (1959), 631–648, is that the generating function $Z(x)$ (known as the *zeta function* of the algebraic variety defined by the equations) is rational. A nice exposition of this result appears in N. Koblitz, *p-adic Numbers, p-adic Analysis, and Zeta-Functions*, second ed., Springer-Verlag, New York/Heidelberg/Berlin, 1984 (Ch. V). For further information on zeta functions, see e.g. R. Hartshorne, *Algebraic Geometry*, Springer-Verlag, New York/Heidelberg/Berlin, 1977 (App. C).

c. See A. A. Jagers and I. Gessel (independently), Solution to E2993, *American Math. Monthly* **93** (1986), 483–484.

5.3. **a.** Since $1 \cdot 3 \cdot 5 \cdots (2n-1) = (2n)!/2^n n!$, we have

$$E_f(x) = \sum_{n\geq 0} 2^{-n} \binom{2n}{n} x^n = (1-2x)^{-1/2},$$

by Exercise 1.4(a). Hence

$$E_f(x)^2 = (1-2x)^{-1} = \sum_{n\geq 0} 2^n n! \frac{x^n}{n!} = E_g(x).$$

b. *First Proof.* $f(n)$ is the number of 1-factors (i.e., graphs whose components are all single edges) on $2n$ vertices, while $g(n)$ is the number of permutations π of $[n]$ with each element of $[n]$ labeled $+$ or $-$. Hence given a labeled permutation π, we want to construct a pair (G, H), where G is a 1-factor on a set of $2k$ vertices labeled by i and i', where i ranges over some subset S of $[n]$, while H is a 1-factor on the $2(n-k)$ complementary vertices j and j', where $j \in T = [n] - S$. Define S (respectively, T) to consist of all i such that the cycle of π containing i has least element labeled $+$ (respectively, $-$). If $\pi(a) = b$, then draw an edge from either a or a' to either b or b', as follows: If a is the least element of its cycle and $a \neq b$, then draw an edge

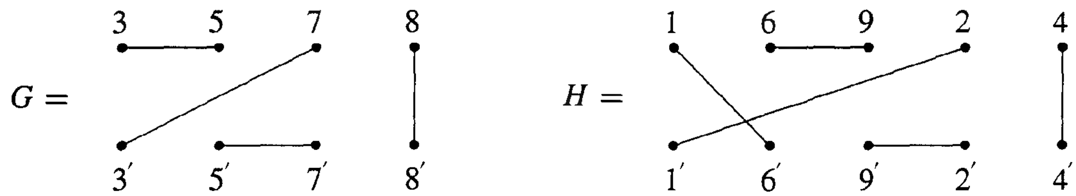


Figure 5-23. A pair of 1-factors.

from a to b (respectively, b') if b is labeled $+$ (respectively, $-$). If neither a nor b is the least element of its cycle, then inductively assume that an edge is incident to either a or a'. Draw a new edge from the vertex a or a' without an edge to b (respectively b') if b is labeled $+$ (respectively $-$). Finally, if b is the least element of its cycle, then only two vertices remain for the last edge – it goes from a or a' (whichever has no edge) to b'. This procedure recursively defines G and H. □

Example. Let

$$\pi = (\underset{-}{1}\ \underset{-}{6}\ \underset{+}{9}\ \underset{-}{2})(\underset{+}{3}\ \underset{+}{5}\ \underset{-}{7})(\underset{-}{4})(\underset{+}{8})$$

Then G and H are given by Figure 5-23.

This bijection is based on work of D. Dumont, *Univ. Beograd. Publ. Elektrotechn. Fak. Ser. Mat. Fiz.*, nos. 634–677 (1979), pp. 116–125 (Prop. 3).

Second Proof. (I. Gessel) It is easy to see that the number of permutations $a_1a_2\cdots a_{2n}$ of the multiset $\{1^2, 2^2, \ldots, n^2\}$ with no subsequence of the form bab with $a < b$ is equal to $f(n)$. (Write down two 1's in one way, then two consecutive 2's in three ways relative to the 1's, then two consecutive 3's in five ways relative to the 1's and 2's, etc.) Hence by Proposition 5.1.1, $E_f(x)^2$ is the exponential generating function for pairs (π, σ), where π is a permutation of some multiset $M = \{i_1^2, \ldots, i_k^2\} \subseteq \{1^2, 2^2, \ldots, n^2\}$ and σ is a permutation of $\{1^2, 2^2, \ldots, n^2\} - M$; and where both π and σ satisfy the above condition on subsequences bab. But to obtain π and σ we can place the two 1's in two ways (i.e., in either π or σ), then the two 2's in four ways, etc., for a total of $2\cdot 4\cdots(2n) = 2^n n!$ ways. □

5.4. a. To obtain a threshold graph G on $[n]$, choose a subset I of $[n]$ to be the set of isolated vertices of G, and choose a threshold graph without isolated vertices on $[n] - I$. From Proposition 5.1.1 it follows that $T(x) = e^x S(x)$.

A threshold graph G with $n \geq 2$ vertices has no isolated vertices if and only if the complement $\bar{G}$ has an isolated vertex. Hence $t(n) = 2s(n)$ for $n \geq 2$. Since $t(0) = s(0) = 1$, $t(1) = 1$, $s(1) = 0$, it follows that $T(x) = 2S(x) + x - 1$.

These results, as well as others related to the enumeration of labeled threshold graphs, are essentially due to J. S. Beissinger and U. N. Peled, *Graphs*

and Combinatorics **3** (1987), 213–219. For further information on threshold graphs, see N. V. R. Mahadev and U. N. Peled, *Threshold Graphs and Related Topics*, Ann. of Discrete Math. **56**, North-Holland, Amsterdam, 1995.

b. Let G be a threshold graph on $[n]$ with no isolated vertices. Define an ordered partition $(B_1, \ldots, B_k)$ of $[n]$ as follows. Let B_1 be the set of isolated vertices of $\bar{G}$, so $\bar{G} = B_1 + G_1$, where G_1 is threshold graph with no isolated vertices. Let B_2 be the set of isolated vertices of $\bar{G}_1$. Iterate this procedure until reaching $\bar{G}_{k-1} = B_k$. We obtain in this way every ordered partition $(B_1, \ldots, B_k)$ of $[n]$ satisfying $\#B_k > 1$. Since there are clearly $nc(n-1)$ ordered partitions $(B_1, \ldots, B_k)$ of $[n]$ satisfying $\#B_k = 1$, it follows that $s(n) = c(n) - nc(n-1)$.

c. The polytope $\mathcal{P}$ of Exercise 4.31 is called a *zonotope* with generators $v_1, \ldots, v_k$. Let Z_n be the zonotope generated by all vectors $e_i + e_j$, $1 \le i < j \le n$, where e_i is the ith unit coordinate vector in $\mathbb{R}^n$. The zonotope Z_n is called the *polytope of degree sequences*. By a well-known duality between hyperplane arrangements and zonotopes (see e.g. A. Björner, M. Las Vergnas, B. Sturmfels, N. White, and G. M. Ziegler, *Oriented Matroids*, Cambridge University Press, Cambridge, 1993 (Prop. 2.2.2)), the number of regions of $\mathcal{T}_n$ is equal to the number of vertices of Z_n. The number of vertices of Z_n was computed by J. S. Beissinger and U. N. Peled, *Graphs and Combinatorics* **3** (1987), 213–219. Further properties of Z_n appear in R. Stanley, in *Applied Geometry and Discrete Mathematics: The Victor Klee Festschrift*, DIMACS Series in Discrete Math. and Theor. Comput. Sci. **4**, American Mathematical Society, 1991, pp. 555–570.

d. This result can be deduced from Exercise 5.50(b). It is also a consequence of the theory of signed graph colorings developed by T. Zaslavsky in *Discrete Math.* **39** (1982), 215–228, and **42** (1982), 287–312 (esp. §5). Is there a more direct proof?

5.5. Let $c_k(n)$ be the number of ways to choose a connected bipartite graph on $[n]$ with k edges. Let $f_k(n)$ (respectively, $g_k(n)$) be the number of ways to choose a weak ordered partition (A, B) of $[n]$ into two parts, and then choose a bipartite graph (respectively, connected bipartite graph) with k edges on $[n]$ such that every edge goes from A to B. Let

$$B(x) = \sum_{n \ge 0} \sum_{k \ge 0} b_k(n) q^k \frac{x^n}{n!},$$

and similarly for $C(x)$, $F(x)$, and $G(x)$. (The sums for $C(x)$ and $G(x)$ start at $n = 1$.) It is easy to see that

$$F(x) = \sum_{n \ge 0} \left(\sum_{i=0}^{n} (1+q)^{i(n-i)} \binom{n}{i} \right) \frac{x^n}{n!}$$

$$F(x) = \exp G(x), \qquad B(x) = \exp C(x), \qquad G(x) = 2C(x),$$

and the proof follows.

5.6. It suffices to assume that $q \in \mathbb{P}$. Let K_{mn} have vertex bipartition (A, B). By an obvious extension of Proposition 5.1.3, the coefficient of $x^m y^n / m!n!$ in

$(e^x + e^y - 1)^q$ is the number of q-tuples $\pi = (S_1, \ldots, S_q)$ where each S_i is a (possibly empty) subset of A or of B, the S_i's are pairwise disjoint, and $\bigcup S_i = A \cup B$. Color the vertices in S_i with the color i. This yields a bijection with the q-tuples π and the q-colorings of K_{mn}, and the proof follows. Note that there is a straightforward extension of this result to the complete multipartite graph $K_{n_1,\ldots,n_k}$, yielding the formula

$$\sum_{n_1,\ldots,n_k \geq 0} \chi(K_{n_1,\ldots,n_k}, q)\frac{x_1^{n_1}}{n_1!}\cdots\frac{x_k^{n_k}}{n_k!} = (e^{x_1} + \cdots + e^{x_k} - k + 1)^q.$$

5.7. **a.** Let $\mathcal{A}_n$ (respectively, $\mathcal{B}_n$) be the set of all pairs (π, σ) such that π and σ are alternating permutations of some complementary subsets S and $[2n] - S$ of $[2n]$ of odd (respectively, even) cardinality. Proposition 5.1.1 shows that the identity $1 + \tan^2 x = \sec^2 x$ follows from giving a bijection from $\mathcal{A}_n$ to $\mathcal{B}_n$ for $n \geq 1$. Suppose that $\pi = a_1a_2\cdots a_k$ and $\sigma = b_1b_2\cdots b_{2n-k}$. Then exactly one of the pairs $(a_1a_2\cdots a_kb_{2n-k}, b_1b_2\cdots b_{2n-k-1})$ and $(a_1a_2\cdots a_{k-1}, b_1b_2\cdots b_{2n}a_k)$ belongs to $\mathcal{B}_n$, and this establishes the desired bijection.

b. The identity (5.87) is equivalent to

$$\sum_{\substack{m,n\geq 0\\ m+n \text{ odd}}} E_{m+n}\frac{x^m}{m!}\frac{y^n}{n!} = \sum_{k\geq 0}[(\tan^k x)(\tan^{k+1} y) + (\tan^{k+1} x)(\tan^k y)].$$

Let $m, n \geq 0$ with $m + n$ odd, and let π be an alternating permutation of $[m + n]$. Then either $n = 0$, or else π can be uniquely factored (as a word $a_1a_2\cdots a_{m+n}$) in the form

$$\pi = e_1\bar{o}_1o_1\bar{o}_2o_2\cdots\bar{o}_ko_k\bar{o}_{k+1}e_2,$$

where (i) e_1 is an alternating permutation (possibly empty) of some subset of $[m]$ of even cardinality, (ii) e_2 is a reverse alternating permutation (possibly empty) of some subset of $[m]$ of even cardinality, (iii) each o_i is a reverse alternating permutation of some subset of $[m]$ of odd cardinality, and (iv) each $\bar{o}_i$ is an alternating permutation of some subset of $[m + 1, m + n]$ of odd cardinality. Using the bijection of (a) (after reversing e_2), we can transform the pair (e_1, e_2) into a pair (e_1', e_2') where the e_i''s are alternating permutations of sets of odd cardinality, unless both e_1 and e_2 are empty. It follows that

$$\begin{aligned}\sum_{\substack{m,n\geq 0\\ m+n \text{ odd}}} E_{m+n}\frac{x^m}{m!}\frac{y^n}{n!} &= \tan x + \sum_{k\geq 0}(1 + \tan^2 x)(\tan^k x)(\tan^{k+1} y)\\ &= \sum_{k\geq 0}[(\tan^k x)(\tan^{k+1} y) + (\tan^{k+1} x)(\tan^k y)].\end{aligned}$$

5.8. For further information on central factorial numbers (where our $T(n, k)$ is denoted $T(2n, 2k)$), see J. Riordan, *Combinatorial Identities*, John Wiley & Sons, New York, 1968 (§6.5). Part (e) is equivalent to a conjecture of J. M. Gandhi, *Amer. Math. Monthly* **77** (1970), 505–506. This conjecture was proved by L. Carlitz, *K. Norske Vidensk. Selsk. Sk.* **9** (1972), 1–4, and by J. Riordan and

P. R. Stein, *Discrete Math.* **5** (1973), 381–388. A combinatorial proof of Gandhi's conjecture was given by J. Françon and G. Viennot, *Discrete Math.* **28** (1979), 21–35. The basic combinatorial property (f)(i) of Genocchi numbers is due to D. Dumont, *Discrete Math.* **1** (1972), 321–327, and *Duke Math. J.* **41** (1974), 305–318. For many further properties of Genocchi numbers, see the survey by G. Viennot, *Séminaire de Théorie des Nombres*, 1981/82, Exp. No. 11, 94 pp., Univ. Bordeaux I, Talence, 1982. A more recent reference is D. Dumont and A. Randrianarivony, *Discrete Math.* **132** (1994), 37–49.

5.9. a. If $C(x)$ is the exponential generating function for the number of connected structures on an n-set, then Corollary 5.1.6 asserts that $F(x) = \exp C(x)$. Hence

$$E_g(x) = \exp \tfrac{1}{2}[C(x) + C(-x)] = \sqrt{F(x)F(-x)}.$$

b. Let $c_k(n)$ be the number of k-component structures that can be put on an n-set, and let

$$F(x,t) = \sum_{n\geq 0}\sum_{k\geq 0} c_k(n)t^k\frac{x^n}{n!}.$$

By Example 5.2.2 we have $F(x,t) = F(x)^t$, so

$$E_e(x) = \frac{1}{2}[F(x,1) + F(x,-1)] = \frac{1}{2}\left(F(x) + \frac{1}{F(x)}\right).$$

This formula was first noted by H. S. Wilf (private communication, 1997).

5.10. a. Put

$$t_i = \begin{cases} 1 & \text{if } k \mid i \\ 0 & \text{if } k \nmid i \end{cases}$$

in (5.30) to get

$$\begin{aligned}\sum_{n\geq 0} f_k(n)\frac{x^n}{n!} &= \exp\sum_{i\geq 1}\frac{x^{ki}}{ki}\\ &= \exp\frac{1}{k}\log(1-x^k)^{-1}\\ &= (1-x^k)^{-1/k}\\ &= \sum_{n\geq 0}\binom{-1/k}{n}(-1)^n x^{kn}.\end{aligned}$$

Hence $f_k(kn) = (kn)!\binom{-1/k}{n}(-1)^n$, which simplifies to the stated answer.

b. Suppose $k \mid n$. We have $n-1$ choices for $\pi(1)$, then $n-2$ choices for $\pi^2(1)$, down to $n-k+1$ choices for $\pi^{k-1}(1)$. For $\pi^k(1)$ we have $n-k+1$ choices, since $\pi^k(1) = 1$ is possible. If $\pi^k(1) \neq 1$ we have $n-k-1$ choices for $\pi^{k+1}(1)$, while if $\pi^k(1) = 1$ we again have $n-k-1$ choices for $\pi(i)$, where i is the least element of $[n]$ not in the cycle $(1, \pi(1), \ldots, \pi^{k-1}(1))$. Continuing this line of reasoning, for our j-th choice we always have $n-j$ possibilities if $k \nmid j$ and $n-j+1$ possibilities if $k \mid j$, yielding the stated answer.

c. Put

$$t_i = \begin{cases} 0 & \text{if } k \mid i \\ 1 & \text{if } k \nmid i \end{cases}$$

in (5.30) to get

$$\begin{aligned}
\sum_{n \geq 0} g_k(n) \frac{x^n}{n!} &= \exp\left(\sum_{i \geq 1} \frac{x^i}{i} - \sum_{i \geq 1} \frac{x^{ki}}{ki} \right) \\
&= \exp\left[\log(1-x)^{-1} - \frac{1}{k} \log(1-x^k)^{-1} \right] \\
&= (1-x^k)^{1/k}(1-x)^{-1} \\
&= (1+x+\cdots+x^{k-1})(1-x^k)^{\frac{1-k}{k}} \\
&= (1+x+\cdots+x^{k-1}) \sum_{n \geq 0} \binom{\frac{1-k}{k}}{n} (-1)^n x^{kn},
\end{aligned}$$

etc. (Compare Exercise 1.44(a).) Note that

$$\left(\sum_{n \geq 0} f_k(n) \frac{x^n}{n!} \right) \left(\sum_{n \geq 0} g_k(n) \frac{x^n}{n!} \right) = \sum_{n \geq 0} n! \frac{x^n}{n!} = \frac{1}{1-x},$$

since every cycle of a permutation either has length divisible by k or length not divisible by k.

d. See E. D. Bolker and A. M. Gleason, *J. Combinatorial Theory (A)* **29** (1980), 236–242, and E. A. Bertram and B. Gordon, *Europ. J. Combinatorics* **10** (1989), 221–226. A combinatorial proof of a generalization of the case $k = 2$ different from (c) appears in R. P. Lewis and S. P. Norton, *Discrete Math.* **138** (1995), 315–318.

5.11. a. A permutation $w \in \mathfrak{S}_n$ has a square root if and only if the number of cycles of each even length $2i$ is even. A simple variant of Example 5.2.10 yields

$$\begin{aligned}
\sum_{n \geq 0} a(n) \frac{x^n}{n!} &= e^x \left(\cosh \frac{x^2}{2} \right) e^{x^3/3} \left(\cosh \frac{x^4}{4} \right) e^{x^5/5} \cdots \\
&= \left(\frac{1+x}{1-x} \right)^{1/2} \prod_{k \geq 1} \cosh \frac{x^{2k}}{2k}.
\end{aligned}$$

This result appears in J. Blum, *J. Combinatorial Theory (A)* **17** (1974), 156–161 (eq. (5)), and [**1.3**, §9.2] (but in this latter reference with the factors $\cosh(x^{2k}/2k)$ misstated as $\cosh(x^{2k}/k)$). These authors are concerned with the asymptotic properties of $a(n)$.

b. Let $F(x) = \sum_n a(n)x^n/n!$. Then by (a) $F(x)/(1+x)$ is even, and the result follows. See J. Blum, *ibid.* (Thm. 1).

5.12. Let $a = u$ and $b = uv^{-1}$. Then $u = a$ and $v = b^{-1}a$, so a and b range over $\mathfrak{S}_n$ as u and v do. Note that $u^2v^{-2} = (aba^{-1})b$. Since every element of $\mathfrak{S}_n$ is conjugate to its inverse, the multiset of elements $aba^{-1}b$ $(a, b \in \mathfrak{S}_n)$ coincides

with the multiset of elements $aba^{-1}b^{-1}$. Thus $f(n)$ is equal to the number of pairs $(a, b) \in \mathfrak{S}_n \times \mathfrak{S}_n$ such that $ab = ba$. (See Exercise 7.69(h).) Since the number $k(a)$ of conjugates of a is the index $[\mathfrak{S}_n : C(a)]$ of the centralizer of a, we have

$$\begin{aligned} f(n) &= \sum_{a \in \mathfrak{S}_n} \#C(a) \\ &= \sum_{a \in \mathfrak{S}_n} \frac{n!}{k(a)} \\ &= n!\, p(n), \end{aligned}$$

where $p(n)$ is the number of partitions of n (the number of conjugacy classes of $\mathfrak{S}_n$). Hence

$$F(x) = \prod_{i \geq 1} (1 - x^i)^{-1}.$$

A less conceptual proof can also be given by considering the possible cycle types of u and v.

Note that the above argument shows the following more general results. First, for any finite group G,

$$\#\{(u, v) \in G \times G : uv = vu\} = k(G) \cdot |G|,$$

where $k(G)$ denotes the number of conjugacy classes in G. (This result was known to P. Erdős and P. Turan, *Acta Math. Hung.* **19** (1968), 413–435 (Thm. IV, proved on p. 431).) Second (using the observation that if $aba^{-1}b = 1$, then b is conjugate to b^{-1}), we have

$$\#\{(u, v) \in G \times G : u^2 = v^2\} = |G| \cdot \iota(G), \tag{5.124}$$

where $\iota(G)$ is the number of "self-inverse" conjugacy classes of G, i.e., conjugacy classes K for which $w \in K \Leftrightarrow w^{-1} \in K$. This result can also be proved using character theory, as done in Exercise 7.69(h) for a situation overlapping the present one when $G = \mathfrak{S}_n$. The problem of computing the left-hand side of (5.124) was posed by R. Stanley, Problem 10654, *Amer. Math. Monthly* **105** (1998), 272.

5.13. **a.** A homomorphism $f : G \to \mathfrak{S}_n$ defines an action of G on $[n]$. The orbits of this action form a partition $\pi \in \Pi_n$. By the exponential formula (Corollary 5.1.6), we have

$$\sum_{n \geq 0} \#\mathrm{Hom}(G, \mathfrak{S}_n) \frac{x^n}{n!} = \exp\left(\sum_d g_d \frac{x^d}{d!} \right),$$

where g_d is the number of *transitive* actions of G on $[d]$. Such an action is obtained by choosing a subgroup H of index d to be the subgroup of G fixing a letter (say 1), and then choosing in $(d-1)!$ ways the letters $1 \neq i \in [d]$ corresponding to the proper cosets of H. Hence $g_d = (d-1)!\, j_d$, and the proof follows.

This result first appeared (though not stated in generating-function form) in I. Dey, *Proc. Glasgow Math. Soc.* **7** (1965), 61–79. The proof given here

appears in K. Wohlfahrt, *Arch. Math.* **29** (1977), 455–457. For some ramifications and generalizations, see T. Müller, *J. London Math. Soc. (2)* **44** (1991), 75–94; *Invent. Math.* **126** (1996), 111–131; and Enumerating representations in finite wreath products, MSRI Preprint No. 1997-048; as well as [**16**, §3.1]. A general survey of the function $j_d(G)$ is given by A. Lubotzky, in *Proceedings of the International Congress of Mathematicians* (Zürich, 1994), vol. 1, Birkhäuser, Basel/Boston/Berlin, 1995, pp. 309–317.

b. If F_s has generators $x_1, \ldots, x_s$, then a homomorphism $\varphi : F_s \to \mathfrak{S}_n$ is determined by any choice of the $\varphi(x_i)$'s. Hence $\#\mathrm{Hom}(F_s, \mathfrak{S}_n) = n!^s$, and the proof follows from (a). A recurrence equivalent to (5.89) was found by M. Hall, Jr., *Canad. J. Math.* **1** (1949), 187–190, and also appears as Theorem 7.2.9 in M. Hall, Jr., *The Theory of Groups*, Macmillan, New York, 1959. Equation (5.89) itself first appeared in [**4.36**, eqn. (21)]. From (5.89) E. Bender, *SIAM Rev.* **16** (1974), 485–515 (§5), has derived an asymptotic expansion for $j_d(F_s)$ for fixed s. For further combinatorial aspects of $j_d(F_2)$, see A. W. M. Dress and R. Franz, *Bayreuth Math. Schr.*, No. 20 (1985), 1–8, and T. Sillke, in *Séminaire Lotharingien de Combinatoire (Oberfranken, 1990)*, Publ. Inst. Rech. Math. Av. **413**, Univ. Louis Pasteur, Strasbourg, 1990, pp. 111–119.

c. Let $m \mid d$. Choose a subgroup H of G of index m, and let $N(H)$ denote its normalizer. Choose an element $z \in N(H)/H$. Define a subgroup K of $G \times \mathbb{Z}$ by

$$K = \{(w, da/m) \in G \times \mathbb{Z} : w \in N(H),\ w = z^a \text{ in } N(H)/H\}.$$

Then $[G \times \mathbb{Z} : K] = d$, and every subgroup K of $G \times \mathbb{Z}$ of index d is obtained uniquely in this way. (This fact is a special case of the description of the subgroups of the direct product of any two groups. See e.g. M. Suzuki, *Group Theory I*, Springer-Verlag, Berlin/Heidelberg/New York, 1982, p. 141, translated from the Japanese edition *Gunron*, Iwanami Shoten, Tokyo, 1977 and 1978.) Once we choose m and H, there are $[N(H) : H]$ choices for z. Since the number of conjugates of H is equal to the index $[G : N(H)]$, we see easily that

$$u_m(G) = \frac{1}{m} \sum_{[G:H]=m} [N(H) : H].$$

It follows that

$$j_d(G \times \mathbb{Z}) = \sum_{m \mid d} m u_m(G), \tag{5.125}$$

and the proof follows from (a) and Exercise 1.40.

NOTE. The numbers $u_d(F_s)$ (where F_s is the free group on s generators) were computed by V. Liskovets, *Dokl. Akad. Nauk BSSR* **15** (1971), 6–9 (in Russian), essentially by using equation (5.90). A messier formula for $u_d(F_s)$ appears in J. H. Kwak and J. Lee, *J. Graph Theory* **23** (1996), 105–109. Note that

$$\#\mathrm{Hom}(\mathbb{Z} \times F_s, \mathfrak{S}_n) = \sum_{u \in \mathfrak{S}_n} (\#C(w))^s,$$

where $C(w)$ denotes the centralizer of w in $\mathfrak{S}_n$ (whose cardinality is given explicitly by equation (7.17)). Using (5.90) it is then not hard to obtain the

formula

$$\prod_{i\geq 1}\left(\sum_{j\geq 0}(j!\,i^j)^{s-1}x^{ij}\right)=\prod_{d\geq 1}(1-x^d)^{-u_d(F_s)},$$

which is equivalent to the formula of Liskovets for $u_d(F_s)$.

d. Observe that

$$c_m(n)=\#\mathrm{Hom}(\mathbb{Z}^m,\mathfrak{S}_n).$$

Now use (c). An equivalent result (stated below) was first proved by J. Bryan and J. Fulman, *Annals of Combinatorics* **2** (1998), 1–6.

NOTE. It is well known (and an easy consequence of Exercise 3.49.5(c) or of equation (5.125)) that

$$\sum_{d\geq 1} j_d(\mathbb{Z}^{m-1})d^{-s}=\zeta(s)\zeta(s-1)\cdots\zeta(s-m+2), \qquad (5.126)$$

where ζ denotes the Riemann zeta function. For the history of this result, see L. Solomon, in *Relations between Combinatorics and Other Parts of Mathematics*, Proc. Symp. Pure Math. **34**, American Mathematical Society, 1979, pp. 309–330. By iterating (5.90) or by using (5.126) directly, we obtain the formula of Bryan and Fulman, viz.,

$$\sum_{n\geq 0} c_m(n)\frac{x^n}{n!}=\prod_{i_1,\ldots,i_{m-1}\geq 1}\left(\frac{1}{1-x^{i_1\cdots i_{m-1}}}\right)^{i_1^{m-2}i_2^{m-3}\cdots i_{m-2}}.$$

e. By (a), we want to show that $h_k(n)=\#\mathrm{Hom}(\mathfrak{S}_k,\mathfrak{S}_n)$. Let $s_m=(m,m+1)\in\mathfrak{S}_k$, $1\leq m\leq k-1$. Given a homomorphism $f:\mathfrak{S}_k\to\mathfrak{S}_n$, define a graph Γ_f on the vertex set $[n]$ by the condition that there is an edge colored m with vertices $a\neq b$ if $f(s_m)(a)=b$. One checks that the conditions (i)–(iii) are equivalent to the well-known Coxeter relations (e.g., J. E. Humphreys, *Reflection Groups and Coxeter Groups*, Cambridge University Press, Cambridge, 1990, §1.9) satisfied by the generators s_m of $\mathfrak{S}_k$.

5.14. a. Take the logarithm of both formulas, subtract one from the other, and solve for y to get

$$1+t^{-1}y=\frac{1-t}{e^{x(t-1)}-t}. \qquad (5.127)$$

Comparing with Exercise 3.81(c) (after correcting a typographical error) shows that the only possible y is as claimed. Since the steps are reversible, the proof follows.

b. While this result can easily be proved using the explicit formula (5.127) and the fact that

$$\frac{d}{dx}\sum_{n\geq 2}A_{n-1}(t)\frac{x^n}{n!}=y,$$

we prefer as usual a combinatorial proof. Define a *connected A-structure* on a finite subset S of $\mathbb{P}$ to consist of a permutation $w=a_1a_2\cdots a_j$ of

S whose smallest element $\min a_i$ is a_1. Define the weight $f(w)$ of w by $f(w) = t^{1+d(w)}$. If $\#S = n$ then it is easy to see that

$$C_n(t) := \sum_w f(w) = \begin{cases} t, & n = 1 \\ A_{n-1}(t), & n > 1, \end{cases}$$

where w ranges over all connected A-structures on S. By the exponential formula (Corollary 5.1.6), we have

$$\exp\left(tx + \sum_{n\geq 2} A_{n-1}(t)\frac{x^n}{n!}\right) = \sum_{n\geq 0} \tilde{A}_n(t)\frac{x^n}{n!},$$

where

$$\tilde{A}_n(t) = \sum_{\pi=\{B_1,\ldots,B_k\}\in\Pi_n} C_{\#B_1}(t)\cdots C_{\#B_k}(t).$$

Given an A-structure w_i on each block B_i of π, where the indexing is chosen so that $\min w_1 > \min w_2 > \cdots > \min w_k$, the concatenation $w = w_1 w_2 \cdots w_k$ is a permutation of $[n]$ such that

$$f(w_1)f(w_2)\cdots f(w_k) = t^{1+d(w)}.$$

Conversely, given $w \in \mathfrak{S}_n$ we can uniquely recover $w_1, w_2, \ldots, w_k$, since the elements $\min w_i$ are the left-to-right minima of w. (Compare the closely related bijection $\mathfrak{S}_n \xrightarrow{\wedge} \mathfrak{S}_n$ of Proposition 1.3.1.) Hence

$$\tilde{A}_n(t) = \sum_{w\in\mathfrak{S}_n} t^{1+d(w)} = A_n(t),$$

completing the proof.

c. It follows from the argument above that the number of left-to-right minima of w is k, the number of blocks of π. The stated formula is now an immediate consequence of the discussion in Example 5.2.2. This result is due to L. Carlitz and R. A. Scoville, *J. Combinatorial Theory* **22** (1977), 129–145 (eqn. (1.13)), with a more computational proof than ours. Carlitz and Scoville state their result in terms of the number of cycles and excedances (which they call "ups") of w, but the bijection $\mathfrak{S}_n \xrightarrow{\wedge} \mathfrak{S}_n$ of Proposition 1.3.1 shows that the two results are equivalent.

d. If a_i is a left-to-right minimum of $w = a_1 a_2 \cdots a_n$, then either $i = 1$ or $i \in D(w)$. Hence $1 + d(w) - m(w) \geq 0$. By (c) we have

$$(1+y)^{q/t} = \sum_{n\geq 0}\left(\sum_{w\in\mathfrak{S}_n} q^{m(w)} t^{1+d(w)-m(w)}\right)\frac{x^n}{n!},$$

and the proof follows.

5.15. **a.**

$$E_f(x) = \exp \sum_{i \geq k} \tfrac{1}{2}(i-1)! \frac{x^i}{i!}$$
$$= (1-x)^{-1/2} \exp\left(-\frac{x}{2} - \frac{x^2}{4} - \cdots - \frac{x^{k-1}}{2(k-1)} \right).$$

b.

$$E_f(x) = \exp\left(\frac{x^2}{2!} + \sum_{i \geq 3} i \frac{x^i}{i!} \right)$$
$$= \exp\left(-x - \frac{x^2}{2} + xe^x \right).$$

c.

$$E_f(x) = \exp\left(\frac{x^4}{4!} + \sum_{i \geq 5} \frac{i(i-2)!}{2} \frac{x^i}{i!} \right)$$
$$= (1-x)^{-x/2} \exp\left(-\tfrac{1}{2}x^2 - \tfrac{1}{4}x^3 - \tfrac{1}{8}x^4\right).$$

d.

$$E_f(x) = \exp\left(x + \sum_{i \geq 2} \frac{i!}{2} \frac{x^i}{i!} \right)$$
$$= \exp\left(\frac{x}{2} + \frac{x}{2(1-x)} \right).$$

5.16. **a.** See R. Stanley, in *Applied Geometry and Discrete Mathematics: The Victor Klee Festschrift*, DIMACS Series in Discrete Math. and Theor. Comput. Sci. **4**, American Mathematical Society, 1991, pp. 555–570 (Cor. 3.4). It would be interesting to have a direct combinatorial proof of (5.91). For some work in this direction, see C. Chan, Ph.D. thesis, M.I.T., 1992 (§3).

b. By Proposition 5.3.2 the number of rooted trees on n vertices is n^{n-1}, with exponential generating function

$$R(x) = \sum_{n \geq 1} n^{n-1} \frac{x^n}{n!}.$$

Hence by Proposition 5.1.3 the exponential generating function for k-tuples of rooted trees is $R(x)^k$, and so for undirected k-cycles of rooted trees (i.e., graphs with exactly one cycle, which is of length $k \geq 3$) is $R(x)^k/2k$.

Let $h(j, n)$ be the number of graphs G on the vertex set $[n]$ such that every component has exactly one cycle, which is of odd length ≥ 3, and such that G has a total of j cycles. (Such graphs have exactly n edges.) Then by the

exponential formula (Cor. 5.1.6) we have

$$\begin{aligned}
\sum_{j,n\geq 0} h(j,n)\frac{t^j x^n}{n!} &= \exp \sum_{k\geq 1} \frac{t}{2(2k+1)} R(x)^{2k+1} \\
&= \exp \frac{t}{2}\left[\frac{1}{2}(\log[1-R(x)]^{-1}\right. \\
&\qquad \left. - \log[1+R(x)]^{-1}) - R(x)\right] \\
&= \left(\frac{1+R(x)}{1-R(x)}\right)^{t/4} e^{-tR(x)/2}.
\end{aligned}$$

Thus,

$$\begin{aligned}
1+\sum_{j,n\geq 1} 2^{j-1}h(j,n)\frac{x^n}{n!} &= \frac{1}{2}\left[\left(\frac{1+R(x)}{1-R(x)}\right)^{1/2} e^{-R(x)}+1\right] \\
&= \frac{1}{2}\left[\left(-1+\frac{2}{1-R(x)}\right)^{1/2} e^{-R(x)}+1\right].
\end{aligned}$$

It is easy to deduce from $R(x) = xe^{R(x)}$ that

$$\frac{1}{1-R(x)} = \sum_{n\geq 0} n^n \frac{x^n}{n!}, \qquad e^{-R(x)} = 1 - \sum_{n\geq 1} (n-1)^{n-1}\frac{x^n}{n!} \tag{5.128}$$

(for the first of these formulas see Exercise 5.42; the second follows from equation (5.67)), so we get

$$\begin{aligned}
1 &+ \sum_{j,n\geq 1} 2^{j-1}h(j,n)\frac{x^n}{n!} \\
&= \frac{1}{2}\left[\left(1+2\sum_{n\geq 1} n^n\frac{x^n}{n!}\right)^{1/2}\left(1-\sum_{n\geq 1}(n-1)^{n-1}\frac{x^n}{n!}\right)+1\right].
\end{aligned}$$

Now by Propositions 5.1.1 and 5.1.6, the exponential generating function for the right-hand side of (5.91) is

$$\left(1+\sum_{j,n\geq 1} 2^{j-1}h(j,n)\frac{x^n}{n!}\right)\cdot e^{T(x)},$$

where $T(x) = \sum_{n\geq 1} n^{n-2}(x^n/n!)$ is the exponential generating function for free trees on the vertex set $[n]$, and the proof follows.

This result appears in R. Stanley, *ibid.*, Cor. 3.6.

5.17. a. Line up all n persons in $n!$ ways. Break the line in $k-1$ of the $n-1$ places between two consecutive persons, in $\binom{n-1}{k-1}$ ways. This yields k lines, but the same k lines could have been obtained in any order, so we must divide by the $k!$ ways of ordering k lines. Thus there are $\frac{n!}{k!}\binom{n-1}{k-1}$ ways. (Exercise 1.11(b) is essentially the same as this one.)

b. Put $f(n)=n!$ and $g(k)=x^k$ in Theorem 5.1.4 (or $f(n)=xn!$ in Corollary 5.1.6).

c. We have

$$\left[\frac{u^r}{r!}\right]\frac{u}{(1-u)^a}=r!\left(\!\!\binom{r}{a-1}\!\!\right),$$

the number of ways to linearly order an r-element set, say $z_1, z_2, \ldots, z_r$, and then to place $a-1$ bars in the spaces between the z_i's or before z_1 (but not after z_r), allowing any number of bars in each space. On the other hand, we have

$$\binom{n+(a-1)k-1}{n-k}=\left(\!\!\binom{n-k+1}{ak-1}\!\!\right),$$

the number of ways to place $ak-1$ bars $B_1, \ldots, B_{ak-1}$ (from left to right) in the spaces between a line of $n-k$ dots, or at the beginning and end of the line, allowing any number of bars in each space. Put a new bar B_0 at the beginning and a new bar B_{ka} at the end. Put a new dot just before the bar B_{ja} for $1 \le j \le k$. We now have n dots in all. Replace them with a permutation of $[n]$ in $n!$ ways. By considering the configuration between $B_{(j-1)a}$ and B_{ja} for $1 \le j \le k$, we see that our structure is equivalent to an *ordered* partition of $[n]$ into k blocks, such that each block has a linear ordering $z_1, \ldots, z_r$ together with $a-1$ bars in the spaces between the z_i's, allowing bars before z_1 but not after z_r. Since there are $k!$ ways of "unordering" the k blocks, equation (5.92) follows from Corollary 5.1.6 (the exponential formula). Equation (5.93) is proved similarly.

Essentially the same argument was found by C. A. Athanasiadis, H. Cohn, and L. W. Shapiro (independently). These identities are also easy to prove algebraically. For instance,

$$\exp\frac{xu}{(1-u)^a}=\sum_{k\ge 0}\frac{u^k}{(1-u)^{ak}}\frac{x^k}{k!}$$

$$=\sum_{k\ge 0}\left(\sum_{n\ge k}\left(\!\!\binom{ak}{n-k}\!\!\right)u^n\right)\frac{x^k}{k!},$$

etc.

d. Choose an $(n-k)$-subset T of $[n]$ in $\binom{n}{k}$ ways. Choose an injection $g: T\to [n]\cup A$ in $(\alpha+n)_{n-k}$ ways. We have $\binom{n}{k}(\alpha+n)_{n-k}$ ways of choosing in all. If $i\in[n]-T$ and i is not in the image of g, then define $\{i\}$ to be a block of π (which of course has a unique linear ordering). If $i\in[n]-T$ and $i=g(j)$ for some j, then there is a unique $m\in T$ for which $g^r(m)=i$ for some $r\ge 1$, and m is not in the image of g. Define a linearly ordered block of π by

$$m > g(m) > g^2(m) > \cdots > g^r(m)=i.$$

The remaining elements of $[n]$ (those not in some block of π) form the set $\bar{S}$, and the restriction of g to $\bar{S}$ defines f.

e. Note that

$$(1-u)^{-\alpha-1} = \sum_{j\geq 0} (\alpha+j)_j \frac{u^j}{j!},$$

and that $(\alpha+j)_j$ is the number of injections $f : \bar{S} \to \bar{S} \cup A$, where $\#\bar{S} = j$. Now use Proposition 5.1.1 and (b). There is also an easy algebraic proof analogous to that given at the end of (c).

The polynomials

$$L_n^{(\alpha)}(x) = \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} (\alpha+n)_{n-k}(-x)^k$$

are the *Laguerre polynomials*. The combinatorial approach used here is due to D. Foata and V. Strehl, in *Enumeration and Design* (D. M. Jackson and S. A. Vanstone, eds.), Academic Press, Toronto/Orlando (1984), pp. 123–140. They derive many additional properties of Laguerre polynomials by similar combinatorial reasoning. Combinatorial approaches toward other classical sequences of polynomials have been undertaken by a number of researchers; see for example G. X. Viennot, *Une Théorie Combinatoire des Polynômes Orthogonaux*, lecture notes, Université de Québec à Montréal, Dépt. de Maths., 1984, 215 pp.; various papers in Springer Lecture Notes in Math., Vol. 1171, Springer-Verlag, Berlin, 1985 (especially pp. 111–157); J. Labelle and Y. N. Yeh, *Studies in Applied Math.* **80** (1989), 25–36. See also Exercise 5.19 for a further example of this type of reasoning. Additional references appear in [**2**, p. xiv].

5.18. If C is a cycle of length n, then the number of distinct cycles that are powers of C is $\phi(n)$ (since the distinct cycles are C^j where $1 \leq j \leq n$ and $(j,n) = 1$). Hence if π has cycles $C_1, C_2, \ldots, C_k$, then the number of permutations equivalent to π is $\prod_{i=1}^{k} \phi(\#C_i)$. Therefore

$$e(n) = \sum_{\pi \in \mathfrak{S}_n} \left(\prod_i \phi(\#C_i)^{-1} \right),$$

where the C_i's are the cycles of π. Now use Corollary 5.1.9.

This result was proposed as a problem by R. Stanley, *Amer. Math. Monthly* **80** (1973), 949, and a solution was given by A. Nijenhius, **82** (1975), 86–87.

5.19. We have

$$K_n(a)K_n(b) = \sum_{\pi,\sigma} a^{c_1(\pi)} b^{c_1(\sigma)},$$

summed over pairs (π, σ) of involutions in $\mathfrak{S}_n$. Represent (π, σ) by a graph $G(\pi, \sigma)$ on the vertex set $[n]$ by putting a red (respectively, blue) edge between i and j if (i, j) is a cycle of π (respectively, σ). If $\pi(i) = i$ (respectively, $\sigma(i) = i$), then we put a red (respectively, blue) loop on the vertex i. (Thus if $\pi(i) = i$ and

$\sigma(i)=i$, then there are two loops on i, one red and one blue.) There are three types of components of $G(\pi, \sigma)$:

(i) A path with a loop at each end and with $2k+1 \geq 1$ vertices, with red and blue edges alternating. There are $(2k+1)!$ such paths, and all have one red and one blue loop. Thus each contribute a factor ab to the term $a^{c_1(\pi)}b^{c_1(\sigma)}$.

(ii) A path as in (i) with $2k \geq 2$ vertices. There are $\frac{1}{2}(2k)!$ paths before we color the edges. One coloring produces two red loops and the other two blue loops, thus contributing a^2 and b^2, respectively, to $a^{c_1(\pi)}b^{c_1(\sigma)}$.

(iii) A cycle of length $2k \geq 2$ with red and blue edges alternating. There are $(2k-1)!$ such cycles, and all have no loops. Thus a cycle contributes a factor of 1 to $a^{c_1(\pi)}b^{c_1(\sigma)}$.

It follows from Corollary 5.1.6 (the exponential formula) that

$$\sum_{n\geq 0} K_n(a)K_n(b)\frac{x^n}{n!} = \exp\left[ab\sum_{k\geq 0}\frac{(2k+1)!x^{2k+1}}{(2k+1)!}\right.$$
$$\left. + \frac{1}{2}(a^2+b^2)\sum_{k\geq 1}\frac{(2k)!x^{2k}}{(2k)!} + \sum_{k\geq 1}\frac{(2k-1)!x^{2k}}{(2k)!}\right]$$
$$= (1-x^2)^{-1/2}\exp\left[\frac{abx+\frac{1}{2}(a^2+b^2)x^2}{1-x^2}\right].$$

The *Hermite polynomials* $H_n(a)$ may be defined by

$$1+\sum_{n\geq 1} H_n(a)\frac{x^n}{n!} = \exp(2ax - x^2). \tag{5.129}$$

(Sometimes a different normalization is used, so the right-hand side of (5.129) becomes $\exp(ax - x^2/2)$.) In terms of the Hermite polynomials, the identity (5.95) becomes

$$\sum_{n\geq 0} H_n(a)H_n(b)\frac{x^n}{n!} = (1-4x^2)^{-1/2}\exp\left[\frac{4abx - 4(a^2+b^2)x^2}{1-4x^2}\right].$$

This identity is known as *Mehler's formula.* M. Schützenberger suggested finding a combinatorial proof, and essentially the above proof was given by D. Foata, *J. Combinatorial Theory (A)* **24** (1978), 367–376. For further results along these lines, see D. Foata, *Advances in Applied Math.* **2** (1981), 250–259, and D. Foata and A. M. Garsia, in *Proc. Symp. Pure Math.* (D. K. Ray-Chaudhuri, ed.), vol. 34, American Mathematical Society, Providence, 1979, pp. 163–179.

5.20. **a.** We want to interpret $xe^{B'(F(x))}$ as the exponential generating function (e.f.g.) for rooted $\mathcal{B}$-graphs on n vertices. By (5.20), $B'(x)$ is the e.f.g. for blocks on an $(n+1)$-element vertex set which are isomorphic to a block in $\mathcal{B}$. Thus by Theorem 5.1.4, $B'(F(x))$ is the e.f.g. for the following structure on an n-element vertex set V. Partition V, and then place a rooted $\mathcal{B}$-graph on each block. Add a new vertex v_0, and place on the set of root vertices together with v_0 a block in $\mathcal{B}$. This is equivalent to a $\mathcal{B}$-graph G on $n+1$ vertices, rooted at a vertex v_0 with the property that only one block of G contains v_0.

It follows from Corollary 5.1.6 that $e^{B'(F(x))}$ is the e.f.g. for the following structure on an n-set V. Choose a partition π of V. Add a root vertex v_A to each block A of π. Place on each set $A \cup \{v_A\}$ a $\mathcal{B}$-graph G_A such that v_A is contained in a single block.

If we identify all the vertices v_A to a single vertex v_*, then we obtain simply a $\mathcal{B}$-graph G on $V \cup \{v_*\}$. Moreover, given G we can uniquely recover the partition π and the graphs G_A by removing v_* from G, seeing the connected components which remain (whose vertex sets will be the A's), and adjoining v_A to each component connected in the same way that v_* was connected to that component. Thus $e^{B'(F(x))}$ is the e.f.g. for connected $\mathcal{B}$-graphs on $V \cup \{v_*\}$, where $\#V = n$.

Lastly it follows from (5.19) that $xe^{B'(F(x))}$ is the e.f.g. for the following structure on an n-set W. Choose an element $w \in W$, then add an element w_* to $W - \{w\}$ and place a connected $\mathcal{B}$-graph on $(W - \{w\}) \cup \{w_*\}$. This is equivalent to rooting W at w and placing a connected $\mathcal{B}$-graph on W. In other words, $xe^{B'(F(x))}$ is the e.f.g. for rooted connected $\mathcal{B}$-graphs on n vertices, and hence coincides with $F(x)$. To obtain (5.97), substitute $F^{\langle -1 \rangle}(x)$ for x in (5.96) and solve for $B'(x) = \sum_{n \geq 1} b(n+1)x^n/n!$.

See Figure 5-24 for an example of the decomposition of rooted connected $\mathcal{B}$-graphs described by $xe^{B'(F(x))}$. Equation (5.96) is known as the *block–tree theorem*, and is due to G. W. Ford and G. E. Uhlenbeck, *Proc. Nat. Acad. Sci. U.S.A.* **42** (1956), 122–128 (the case $y_0 = 1$ of (7)). Ford and Uhlenbeck in fact prove a more general result where they keep track of the number of occurrences of each block in a $\mathcal{B}$-graph G. They then use Lagrange inversion to obtain that the number of $\mathcal{B}$-graphs on an n-element vertex set with k_B blocks isomorphic to B is equal to

$$\frac{n! \cdot n^{\Sigma_B k_B - 1}}{\prod_B \left(\frac{|\mathrm{Aut}\, B|}{p_B} \right)^{k_B} k_B!}$$

where the block B has p_B vertices.

b. Let $\mathcal{B}$ be the set of all blocks without multiple edges. A $\mathcal{B}$-graph is just a connected graph without multiple edges. Letting $F(x)$ and $B(x)$ be as in (a),

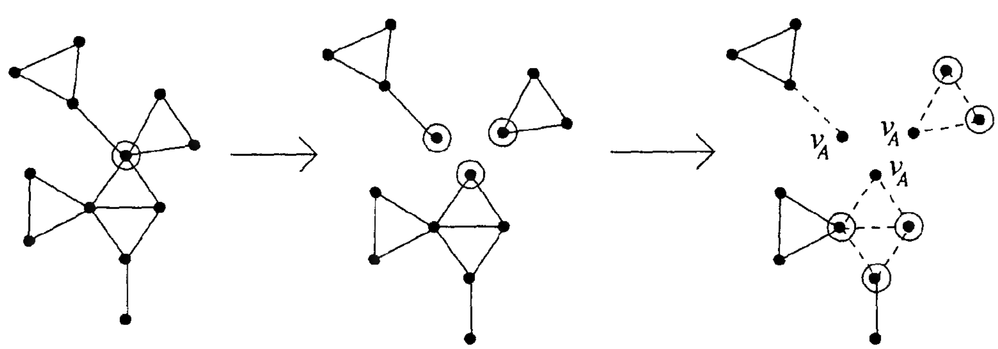


Figure 5-24. The block–tree decomposition.

by (5.37) and (5.21) we have

$$F(x) = x\frac{d}{dx}\log\sum_{n\geq 0} 2^{\binom{n}{2}}\frac{x^n}{n!}.$$

Now use (5.97).

5.21. Let $u = u_1u_2\cdots u_n$, where $u_i \in \mathcal{A}$. Represent u as a row of n dots, and connect two adjacent dots if they belong to the same word of $\mathcal{B}$ when u is factored into words in $\mathcal{B}$. If $\pi = a_1a_2\cdots a_n$, then place a_i below the i-th dot. For instance, if $u = u_1u_2\cdots u_9$ where $u_1 \cdot u_2u_3u_4 \cdot u_5 \cdot u_6u_7 \cdot u_8u_9$ represents the factorization of u into words in $\mathcal{B}$, and if $\pi = 529367148$, then we obtain the diagram

$$\begin{array}{ccccccccccccccccc}
u_1 & & u_2 & & u_3 & & u_4 & & u_5 & & u_6 & & u_7 & & u_8 & & u_9 \\
\bullet & & \bullet & - & \bullet & - & \bullet & & \bullet & & \bullet & - & \bullet & & \bullet & - & \bullet \\
5 & & 2 & & 9 & & 3 & & 6 & & 7 & & 1 & & 4 & & 8
\end{array}$$

Consider the subsequence ρ of π consisting of the labels of the *first* elements of each connected string. For the above example, we get $\rho = 52674$. Draw a bar before all left-to-right maxima (except the first) of the sequence ρ. For $\rho = 52674$, the left-to-right maxima are 5, 6, and 7. Thus we get

$$\begin{array}{ccccccccccccccccc}
u_1 & & u_2 & & u_3 & & u_4 & \vert & u_5 & \vert & u_6 & & u_7 & & u_8 & & u_9 \\
\bullet & & \bullet & - & \bullet & - & \bullet & \vert & \bullet & \vert & \bullet & - & \bullet & & \bullet & - & \bullet \\
5 & & 2 & & 9 & & 3 & \vert & 6 & \vert & 7 & & 1 & & 4 & & 8
\end{array}$$

For each sequence v of u_i's separated by bars, write down the cyclic permutation of v whose first element corresponds to the largest possible element of π. Arrange in a cycle the elements of π which are below v. For our example, we obtain:

$$\begin{array}{rl}
u_3u_4u_1u_2 & (5293) \\
u_5 & (6) \\
u_9u_6u_7u_8 & (7148)
\end{array}$$

We leave it to the reader to verify that this procedure establishes the desired bijection.

This bijection is due to I. Gessel.

5.22. To form a graph with every component a cycle on the vertex set $[n+1]$, first choose such a graph G on the vertex set $[n]$ (in $L(n)$ ways). Then insert the vertex $n+1$ into it, either as an isolated vertex (one way) or by choosing an edge e of G and inserting $n+1$ in the middle of it (n ways). Every allowable graph on $[n+1]$ will arise exactly once, except that the two ways of inserting $n+1$ into a 2-cycle (double edge) result in the same graph. There are $\binom{n}{2}$ possible edges, and $L(n-2)$ graphs which contain a given one of them. Hence $L(n+1) = (n+1)L(n) - \binom{n}{2}L(n-2)$, as desired.

This result was first proved by I. Schur, *Arch. Math. Phys. Series 3* **27** (1918), 162, in a less combinatorial fashion. See also [**53**, Problem VII.45].

5.23. Let N be a cloud. Identify the line δ_i with the node i, and the intersection $\delta_i \cap \delta_j \in N$ with the edge $\{i, j\}$.

This exercise is taken from [**2.3**, pp. 273–277]. The connection between clouds and graphs goes back to W. A. Whitworth, *Choice and Chance*, Bell, 1901

(reprinted by Hafner, 1965), Exer. 160, p. 269. Whitworth erroneously claimed that $c(n) = \frac{1}{2}(n-1)!$. His error was corrected by Robin Robinson, *Amer. Math. Monthly* **58** (1951), 462–469, who obtained the recurrence for $c(n)$ given in Example 5.2.8 (where $T_n^*(2)$ is used instead of $c(n)$) by simple combinatorial reasoning. The generating function (5.29) was derived from the recurrence in an editorial note [**19**], and was used to complete an asymptotic formula for $c(n)$ partially found by Robinson. Some congruence properties of $c(n)$ were later given by L. Carlitz in *Amer. Math. Monthly* **61** (1954), 407–411, and **67** (1960), 961–966.

5.24. a. By Example 4.6.33(b), the vertex set $V(\Sigma_n)$ of Σ_n satisfies

$$V(\Sigma_n) \subseteq \left\{ \tfrac{1}{2}(P + P^t) : \ P \text{ is an } n \times n \text{ permutation matrix} \right\}.$$

It is fairly straightforward to check which matrices $\frac{1}{2}(P + P^t)$ are actually vertices. See M. Katz, *J. Combinatorial Theory* **8** (1970), 417–423 (Thm. 1).

b. Let $\frac{1}{2}(P + P^t) \in V(\Sigma_n)$. Suppose that P corresponds to the permutation π of $[n]$. Define a graph $G = G(P)$ on the vertex set $[n]$ by drawing an edge between i and j if $\pi(i) = j$ or $\pi(j) = i$. By (a), the components of G are single vertices with one loop, single edges, or odd cycles of length ≥ 3. Moreover, every such G corresponds to a unique vertex of Σ_n (though not necessarily to a unique P). There is one way to place a loop on one vertex or an edge on two vertices, and $\frac{1}{2}(2i)!$ ways to place a cycle on $2i + 1 \geq 3$ vertices. Hence

$$\begin{aligned}
\sum_{n\geq 0} M(n)\frac{x^n}{n!} &= \exp\left(x + \frac{x^2}{2} + \sum_{i\geq 1} \frac{1}{2}(2i)! \frac{x^{2i+1}}{(2i+1)!} \right) \\
&= \exp\left(\frac{x}{2} + \frac{x^2}{2} + \frac{1}{2}\sum_{i\geq 0} \frac{x^{2i+1}}{2i+1} \right) \\
&= \exp\left(\frac{x}{2} + \frac{x^2}{2} + \frac{1}{4}\left[\log(1-x)^{-1} - \log(1+x)^{-1} \right] \right) \\
&= \left(\frac{1+x}{1-x} \right)^{1/4} \exp\left(\frac{x}{2} + \frac{x^2}{2} \right).
\end{aligned}$$

An equivalent result (but not stated in terms of generating functions) appears in M. Katz, *ibid.* (Thm. 2).

c. Take the logarithm of (5.98) and differentiate to get

$$\begin{aligned}
\sum_{n\geq 0} M(n+1)\frac{x^n}{n!} &= \left(\sum_{n\geq 0} M(n)\frac{x^n}{n!} \right) \frac{d}{dx} \left(\frac{1}{4}\log(1+x) \right. \\
&\qquad \left. - \frac{1}{4}\log(1-x) + \frac{x}{2} + \frac{x^2}{2} \right) \\
&= \left(\sum_{n\geq 0} M(n)\frac{x^n}{n!} \right) \left(\frac{1}{2(1-x^2)} + \frac{1}{2} + x \right).
\end{aligned}$$

Multiply by $2(1-x^2)$ and take the coefficient of $x^n/n!$ on both sides to obtain

$$M(n+1) = M(n)+n^2M(n-1)-\binom{n}{2}M(n-2)-n(n-1)(n-2)M(n-3).$$

This recurrence first appeared (with a misprint) in [**6.70**, Example 2.8].

5.25. **a.** This result is stated without proof (in a more complicated but equivalent form) by M. Katz, *J. Combinatorial Theory* **8** (1970), 417–423, and proved by the same author in *J. Math. Anal. Appl.* **37** (1972), 576–579.

b. Arguing as in the solution to Exercise 5.24(b), the graph G corresponding to a matrix now can have as a component a single vertex with no loop. (Removing a 1 from the main diagonal converts a loop to a loopless vertex.) Thus when applying the exponential formula as in Exercise 5.24(b), we obtain an additional factor of e^x. (An erroneous generating function appears in [**6.70**, Example 2.8].)

c. As in Exercise 5.24(c), we obtain

$$\sum_{n\geq 0} M^*(n+1)\frac{x^n}{n!} = \left(\sum_{n\geq 0} M^*(n)\frac{x^n}{n!}\right)\left(\frac{1}{2(1-x^2)}+\frac{3}{2}+x\right),$$

from which there follows

$$M^*(n+1) = 2M^*(n)+n^2M^*(n-1)-3\binom{n}{2}M^*(n-2)$$
$$-n(n-1)(n-2)M^*(n-3).$$

Is there a combinatorial proof, analogous to Exercise 5.22?

5.26. Given a set X, let $\mathcal{D}(X)$ denote the set of all subsets S of $2^X - \{\emptyset\}$ such that any two elements of S are either disjoint or comparable. Write $\mathcal{D}(n)$ for $\mathcal{D}([n])$. Since for $n \geq 1$ we have $S \in \mathcal{D}(n)$ and $[n] \notin S$ if and only if $[n] \notin S$ and $S \cup \{[n]\} \in \mathcal{D}(n)$, it follows that $F(x) = 1 + 2G(x)$. Now let $S \in \mathcal{D}(n)$, and regard S as a poset ordered by inclusion. It is not hard to see that S is a disjoint union of rooted trees, with the successors of any vertex being disjoint subsets of $[n]$. Hence S can be uniquely obtained as follows. Choose a partition $\pi = \{B_1, \ldots, B_k\}$ of $[n]$. For each block B_i of π, choose a set $S_i \in \mathcal{D}(B_i)$ such that $B_i \in S_i$ (in $g(\#B_i)$ ways). If $\#B_i = 1$, then we can also choose to have $B_i \notin S_i$. Finally let $S = \bigcup S_i$. Since there are $g(\#B_i)$ choices for each B_i and one extra choice when $\#B_i = 1$, it follows from Corollary 5.1.6 that $F(x) = e^{x+G(x)}$.
This exercise is due to I. Gessel.

5.27. Given an edge-labeled tree T with n edges, choose a vertex of T in $n+1$ ways and label it 0. Then "push" each edge label to the vertex of that edge farthest from 0. We obtain a bijection between (a) the $(n+1)e(n)$ ways to choose T and the vertex 0, and (b) the $(n+1)^{n-1}$ ways to choose a labeled tree on $n+1$ vertices. Hence $e(n) = (n+1)^{n-2}$. Essentially this bijection (though not an explicit statement of the formula $e(n) = (n+1)^{n-2}$) appears in J. Riordan, *Acta Math.* **97** (1957), 211–225 (see equation (17)), though there may be much earlier references.

*** 5.28.** Suppose that the tree T on the vertex set $[n]$ has ordered degree sequence $(d_1, \ldots, d_n)$ (i.e., vertex i has d_i adjacent vertices), where necessarily $\sum d_i = 2n-2$. Choose a vertex of degree one (endpoint), and adjoin vertices one at a time to the graph already constructed, keeping the graph connected. Color each edge as it is added to the graph. For the first edge we have k choices of colors. If one edge of a vertex of degree d has been colored, then there are $(d-1)!\binom{k-1}{d-1}$ ways to color the others. It follows easily from Theorem 5.3.4 that the number of free trees with ordered degree sequence $(d_1, \ldots, d_n)$ is equal to the multinomial coefficient

$$\binom{n-2}{d_1-1, d_2-1, \ldots, d_n-1}.$$

Hence the total number of k-edge colored trees is given by

$$\begin{aligned} T_k(n) &= k(n-2)! \sum_{d_1+\cdots+d_n=2n-2} \prod_{i=1}^{n} \binom{k-1}{d_i-1} \\ &= k(n-2)![x^{n-2}]\left((1+x)^{k-1}\right)^n \\ &= k(n-2)!\binom{(k-1)n}{n-2}. \end{aligned}$$

This result is due to I. Gessel (private communication). Is there a simple bijective proof?

5.29. **a.** If $F \in P_n$ has rank i, then any of the i edges of F can be removed from F to obtain an element that F covers. Hence F covers i elements. To obtain an element that covers F, choose a vertex v of F in n ways, and then choose a connected component T of F not containing v in $n-i-1$ ways. Attach the root of T below v. Thus F is covered by $(n-i-1)n$ elements.

* **b.** Let $M(n)$ denote the number of maximal chains in P_n. We obtain a maximal chain by choosing a maximal element of P_n in $r(n)$ ways, then an element that it covers in $n-1$ ways, etc. Hence $M(n) = r(n)(n-1)!$. On the other hand, we can choose a maximal chain by starting at $\hat{0}$, choosing an element u covering $\hat{0}$ in $(n-1)n$ ways, then an element covering u in $(n-2)n$ ways, etc. Hence $M(n) = n^{n-1}(n-1)!$, so $r(n) = n^{n-1}$.

This elegant proof appears in J. Pitman, Coalescent random forests, *J. Combinatorial Theory (A)*, to appear. The same reasoning can be used to compute the number $p_k(n)$ of planted forests on $[n]$ with k components (i.e., the number of elements of P_n of rank $n-k$), as was done by other methods in the text (Proposition 5.3.2 and Example 5.4.4). Note also that P_n, with a $\hat{1}$ adjoined, is a *triangular poset* in the sense of Exercise 3.79 (except for not having all maximal chains of infinite length).

c. The poset P_n is *simplicial*, i.e., every interval $[\hat{0}, t]$ is isomorphic to a boolean algebra. (In fact, P_n is the face poset of a simplicial complex.) It follows from

Example 3.8.3 and the recurrence (3.14) defining the Möbius function that

$$\mu_n := \mu(\hat{0}, \hat{1}) = -p_n(n) + p_{n-1}(n) - \cdots \pm p_1(n),$$

where $p_k(n)$ denotes the number of planted forests on $[n]$ with k components. If $R(x)$ denotes the exponential generating function for rooted trees (defined in Section 5.3), then by the exponential formula (Corollary 5.1.6) we have

$$\sum_{n\geq 1} \mu_n \frac{x^n}{n!} = 1 - e^{-R(-x)}.$$

Now use the second formula of equation (5.128).

5.30. **First solution**. Linearly order $R \cup S$ by $1 < \cdots < r < 1' < \cdots < s'$. Given T, define a sequence $T_1, T_2, \ldots, T_{r+s-2}$ as follows: set $T_1 = T$. If $i \leq r+s-2$ and T_i has been defined, then define T_{i+1} to be the tree obtained from T_i by removing its *largest endpoint* v_i (and the edge incident to v_i). For each i we also define a pair (u_i, u_i') of sequences (or words) $u_i \in R^*$ and $u_i' \in S^*$ as follows. Set $(u_0, u_0') = (\emptyset, \emptyset)$, where $\emptyset$ denotes the empty word. Let t_i be the unique vertex of T_i adjacent to v_i. If $t_i \in R$ then set $(u_i, u_i') = (u_{i-1}t_i, u_{i-1}')$. If $t_i \in S$ then set $(u_i, u_i') = (u_{i-1}, u_{i-1}'t_i)$. Thus for the tree T we obtain a pair of words $(u, u') = (u_{r+s-2}, u_{r+s-2}')$, where $u_{r+s-2} \in R_{s-1}^*$, $u_{r+s-2}' \in S_{r-1}^*$. As in the first proof of Proposition 5.3.2, the correspondence $T \mapsto (u, u')$ is a bijection between free bipartite trees on (R, S) and the set $R_{s-1}^* \times S_{r-1}^*$. Moreover, a vertex t appears in u and u' one fewer times than its degree, from which (5.100) follows.

Example. For the tree T of Figure 5-25, we have $(u, u') = (3113, 3'1'3')$, and T_7 consists of a single edge connecting 1 and $3'$.

Second solution. There are $r^s s^r$ functions $f : R \cup S \to R \cup S$ satisfying $f(R) \subseteq S$ and $f(S) \subseteq R$. Let D_f denote the digraph of such a function f. The "cyclic part" of D_f corresponds to a permutation π of some subset $R_1 \cup S_1$ of $R \cup S$, where $\pi(R_1) = S_1$ and $\pi(S_1) = R_1$. Linearly order $R_1 \cup S_1$ as $a_1' < a_1 < a_2' < a_2 < \cdots < a_j' < a_j$, where $a_1' < a_2' < \cdots < a_j'$ and $a_1 < a_2 < \cdots < a_j$ as integers. This linear ordering allows π to be written as a word $w = b_1 b_1' b_2 b_2' \cdots b_j b_j'$, where $\pi(a_i') = b_i$, $\pi(a_i) = b_i'$. Regard the word w as a path P in a (bipartite) graph. Circle the endpoints b_1 and b_j'. Attach to each vertex t of P the tree that is attached to t in D_f (with the arrows removed from each edge), yielding a bipartite tree T on (R, S) with a root in R and a root in S. As in the second proof of Proposition 5.3.2, the map $f \mapsto T$ is a bijection

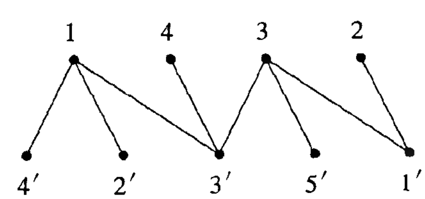


Figure 5-25. A labeled bipartite tree.

between functions $f : R \cup S \to R \cup S$ with $f(R) \subseteq S$ and $f(S) \subseteq R$, and "bi-rooted" bipartite trees on (R, S) with a root in R and a root in S. Moreover, if t is not a root then $\deg_T t = 1 + \#f^{-1}(t)$, while if t is a root then $\deg_T t = \#f^{-1}(t)$. It follows that

$$\begin{aligned}\sum_{\substack{a\in R\\ b'\in S}} (x_a y_b)\left(x_1^{-1}\cdots x_r^{-1}\right)\left(y_1^{-1}\cdots y_s^{-1}\right)\sum_T \left(\prod_{i\in R} x_i^{\deg i}\right)\left(\prod_{j'\in S} y_j^{\deg j'}\right)\\ = (x_1+\cdots+x_r)^s(y_1+\cdots+y_s)^r,\end{aligned} \tag{5.130}$$

where T ranges over all free bipartite trees on (R, S). Then (5.100) follows immediately from (5.130).

Example. Let T be as in Figure 5-25. Suppose we choose 4 and $1'$ as the roots. The corresponding path P is $43'31'$, so the cyclic part of f written in two-line notation is

$$\begin{pmatrix} 1' & 3 & 3' & 4 \\ 4 & 3' & 3 & 1' \end{pmatrix},$$

and in cycle notation is $(1', 4)(3', 3)$. The digraph D_f is shown in Figure 5-26.

The number $c(K_{rs})$ of spanning trees of K_{rs} was first obtained (by different methods than here) by M. Fiedler and J. Sedláček, *Časopis pro Pěstování Matematiky* **83** (1958), 214–225; T. L. Austin, *Canad. J. Math.* **12** (1960), 535–545 (a special case of Thm. II); and H. I. Scoins, *Proc. Camb. Phil. Soc.* **58** (1962), 12–16.

5.31. **a.** Easy.

b. Given a function $f : S \to T$, let D_f be the digraph with vertex set $S \cup T$ and edges $s \to f(s)$ for $s \in S$. Now fix $A \subseteq [n]$, and consider the sum

$$F_A = \sum_g \prod_{i\in A} x_{g(i)}, \tag{5.131}$$

where g ranges over all *acyclic* (i.e., D_g has no directed cycles) functions $g : A \to A \cup \{n+1\}$. Then D_g is an oriented tree with root $n+1$, and the exponent of x_j in the product in (5.131) is equal to $(\deg j) - 1$ if $j \neq n+1$, and to $\deg(n+1)$ if $j = n+1$, where $\deg k$ denotes the total number of vertices adjacent to k (ignoring the direction of the edges). Since the root and orientation of D_g can be determined from the underlying free tree on

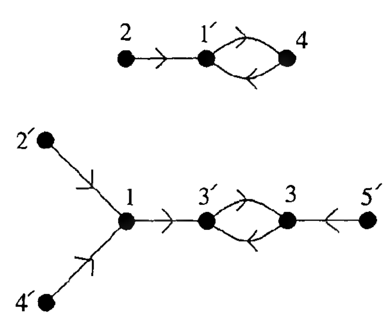


Figure 5-26. The digraph D_f of a function $f : R \cup S \to R \cup S$.

$A \cup \{n+1\}$, it follows from Theorem 5.3.4 that

$$F_A = x_{n+1}\left(x_{n+1} + \sum_{i\in A} x_i\right)^{\#A-1}.$$

Next consider

$$G_A = \sum_h \prod_{i\in A'} x_{h(i)},$$

where h ranges over *all* functions $h : A' \to A' \cup [n+2]$. By (a), we have

$$G_A = \left(x_{n+2} + \sum_{i\in A'} x_i\right)^{n-\#A}.$$

If now $f : [n] \to [n+2]$, then the component of D_f containing $n+1$ will be equal to D_g for a unique $A \subseteq [n]$ and acyclic $g : A \to A \cup \{n+1\}$. The remainder of D_f is equal to D_h for a unique $h : A' \to A' \cup \{n+2\}$. Thus

$$\begin{aligned}(x_1 + \cdots + x_{n+2})^n &= \sum_{f:[n]\to[n+2]} \prod_{i=1}^n x_{f(i)}\\ &= \sum_{A\subseteq[n]} F_A G_A,\end{aligned}$$

and the proof follows.

This result is equivalent to one of A. Hurwitz, *Acta Math.* **26** (1902), 199–203. See also [**2.3**, Exer. 20, p. 163] and [**41**, Exer. 2.3.44-30]. The proof given here is a minor variation of one of J. Françon, *Discrete Math.* **8** (1974), 331–343 (repeated in [**2.3**, pp. 129–130]). Françon uses an elegant "coding" of functions $[n] \to [n]$ due to D. Foata and A. Fuchs, *J. Combinatorial Theory* **8** (1970), 361–375, and obtains many related results in a systematic way. For a generalization, see A. J. Stam, *J. Math. Anal. Appl.* **122** (1987), 439–443.

c. Put $x_{n+1} = x,\ x_{n+2} = y + nz,\ x_1 = x_2 = \cdots = x_n = -z$ and collect the A such that $\#A = k$ in (b). This famous identity, one of several equivalent ones called "Abel's identity" (see the fourth entry of Exercise 5.37(b)), is to due N. Abel, *J. Reine Angew. Math.* (*=Crelle's J.*) **1** (1826), 159–160, or *Oeuvres Complètes*, vol. 1, p. 102. For some other proofs, see [**2.3**, pp. 128–129] and [**41**, Exer. 1.2.6-51]. For additional references, see H. W. Gould, *Amer. Math. Monthly* **69** (1962), 572. For a combinatorial treatment of many identities related to Abel's identity, see V. Strehl, *Discrete Math.* **99** (1992), 321–340.

d. This is equivalent to the case $x = 1,\ y = n,\ z = -1$ of (c). (It can also be proved directly by considering functions $[n] \to [n+1]$.)

5.32. a. Fix $j \in \mathbb{P}$. Given a rooted tree τ, let $w(\tau) = \prod t_{jk}^{a_k}$, where τ has a_k vertices at distance k from the root. By a simple refinement of (5.41), we have

$$\sum_{n\geq 1}\left[\sum_\tau w(\tau)\right]\frac{x^n}{n!} = t_{j0}xe^{t_{j1}xe^{t_{j2}xe^{\cdot^{\cdot^{\cdot}}}}} = E_j, \qquad \text{say,}$$

where τ ranges over all rooted trees on $[n]$.

Now let C be a collection of j such trees $\tau_1, \ldots, \tau_j$ arranged in a j-cycle, and define $w(C) = \prod w(\tau_i)$. Then

$$\sum_{n\geq 1}\left[\sum_C w(C)\right]\frac{x^n}{n!} = \sum_{j\geq 1}\frac{1}{j}E_j^j,$$

where C ranges over all "cycles of rooted trees" on the vertex set $[n]$, since by Proposition 5.1.3, E_j^j enumerates j-tuples $(\tau_1, \ldots, \tau_j)$ of rooted trees, and each j-cycle corresponds to j distinct j-tuples.

Finally by Corollary 5.1.6 the exponential generating function for disjoint unions of cycles of rooted trees on $[n]$ (or digraphs of functions $f : [n] \to [n]$) is given by

$$\exp\sum_{j\geq 1}\frac{1}{j}E_j^j,$$

as desired.

b. $\tilde{Z}_n(t_{jk} = 1)$ is just the number n^n of functions $f : [n] \to [n]$, so the first equality follows. The second equality is a consequence of (5.41) and Proposition 5.3.2.

c. A necessary and sufficient condition that $f^a = f^{a+b}$ is that (i) every cycle of D_f has length dividing b, and (ii) every vertex of D_f is at distance at most a from a cycle. Hence (c) follows by substituting in (a)

$$t_{jk} = \begin{cases} 1 & \text{if } j \mid b \text{ and } k \leq a \\ 0 & \text{otherwise.} \end{cases}$$

d. Since $f = f^{1+b}$ for some $b \in \mathbb{P}$ if and only if every vertex of D_f is at distance at most one from a cycle, we obtain from (a) by setting $t_{j0} = t_{j1} = 1,\ t_{jk} = 0$ if $k > 1$ (or from (c) by letting $b = m!$ and $m \to \infty$) that

$$\begin{aligned}
\sum_{n\geq 0} h(n)\frac{x^n}{n!} &= \exp\sum_{j\geq 1}\frac{1}{j}(xe^x)^j \\
&= \exp\log(1 - xe^x)^{-1} \\
&= (1 - xe^x)^{-1} \\
&= \sum_{m\geq 0} x^m e^{mx} \\
&= \sum_{m\geq 0}\sum_{r\geq 0} m^r\frac{x^{m+r}}{r!} \\
&= \sum_{n\geq 0}\left(\sum_{k=1}^{n} k^{n-k}(n)_k\right)\frac{x^n}{n!},
\end{aligned}$$

so (5.104) follows.

Putting $b = 1$ in (5.103) yields

$$\begin{aligned}\sum_{n\geq 0} g(n)\frac{x^n}{n!} &= \exp xe^x \\ &= \sum_{m\geq 0}\frac{x^m e^{mx}}{m!},\end{aligned} \tag{5.132}$$

and (5.105) follows in a similar manner to (5.104).

e. Note that f satisfies $f^a = f^{a+1}$ for some $a \in \mathbb{P}$ if and only if every cycle of the digraph D_f has length one. Hence we want the number of planted forests on $[n]$, which by Proposition 5.3.2 is $(n+1)^{n-1}$.

f. While a proof using generating functions is certainly possible, there is a very simple direct argument. Namely, for each $i \in [n]$, we have $n-1$ choices for $f(i)$. Hence there are $(n-1)^n$ such functions. Note that the proportion $P(n)$ of functions $f:[n] \to [n]$ without fixed points is $(n-1)^n/n^n$, so $\lim_{n\to\infty} P(n) = 1/e$. From equation (2.12) this is also the limiting value of the proportion of *permutations* $f:[n] \to [n]$ without fixed points.

Equation (5.101) can be deduced from the general "composition theorem" of B. Harris and L. Schoenfeld, in *Graph Theory and Its Applications* (B. Harris, ed.), Academic Press, New York/London, 1970, pp. 215–252. In that paper equation (5.102) is essentially derived, though it is not explicitly written down. Special cases of (5.102) had appeared in earlier papers; in particular, equations (5.105) and (5.132) are obtained by B. Harris and L. Schoenfeld, *J. Combinatorial Theory* **3** (1967), 122–135, along with considerable additional information concerning the number $g(n)$ of idempotents in the symmetric semigroup Λ_n. The first explicit statement of (5.102) seems to be [**3.16**, §3.3.15, Exer. 3.3.31], and a refinement appears in [**16**, §3.2]. For asymptotic properties of Λ_n, see B. Harris, *J. Combinatorial Theory (A)* **15** (1973), 66–74, and B. Harris, *Studies in Pure Mathematics*, Birkhäuser, Basel/Boston/Stuttgart, 1983, pp. 285–290.

5.33. The functions c and $2-\zeta$ are *not* multiplicative, so Theorem 5.1.11 does not apply. Since

$$(2-\zeta)^{-1} = \sum_{k\geq 0}(\zeta-1)^k,$$

the correct generating function is the unappealing

$$E_c(x) = \sum_{k\geq 0} f^{(k)}(x),$$

where $f(x) = f^{(1)}(x) = e^x - x - 1$ (set $f^{(0)}(x) = x$).

5.34. **a.** Straightforward generalization of Theorem 5.1.11.

b. Let $\zeta : \mathbb{P} \to K$ be given by $\zeta(n) = 1$ for all n. Thus $\zeta^2(n) = q_n$ and $\zeta^{-1}(n) = \mu_n$. Since $\varphi(\zeta) = e_k(x)$, the result follows from (a).

c. Define $\zeta_t : \mathbb{P} \to K$ by $\zeta_t(n) = t^n$. Then $\chi_n(t) = \mu\zeta_t(n)$. Now

$$\varphi(\zeta_t) = \sum_{n\geq 0} t^n \frac{x^{kn+1}}{(kn+1)!}$$

$$= t^{-1/k} e_k(t^{1/k}x),$$

while $\varphi(\mu) = e_k^{\langle -1\rangle}(x)$. Thus (5.106) follows from (a). When $k=2$, (5.106) becomes

$$t\sum_{n\geq 0} \chi_n(t^2)\frac{x^{2n+1}}{(2n+1)!} = \sinh(t\sinh^{-1} x).$$

To get (5.107), use Exercise 1.44(c).

For further results on Ψ_n and related posets, see A. R. Calderbank, P. J. Hanlon, and R. W. Robinson, *Proc. London Math. Soc. (3)* **53** (1986), 288–320; S. Sundaram, *Contemporary Math.* **178** (1994), 277–309; and the references given in this latter paper.

5.35. a. Let T be a plane tree with $n+1$ vertices for which s_i internal vertices have i successors. Label the vertices of T in preorder with the numbers $0, 1, \ldots, n$. Let $\pi(T)$ be the partition of $[n]$ whose blocks are the sets of vertices with a common parent. This sets up a bijection with noncrossing partitions of $[n]$ of type $s_1, \ldots, s_n$, and the proof follows from Theorem 5.3.10. This result was first proved (by other means) by G. Kreweras, *Discrete Math.* **1** (1972), 333–350 (Thm. 4). The bijective proof just sketched was first found by P. H. Edelman (unpublished). Later, independently, N. Dershowitz and S. Zaks, *Discrete Math.* **62** (1986), 215–218, gave the same bijection between plane trees amd noncrossing partitions, though they don't explicitly mention enumerating noncrossing partitions by type.

b. Assume char $K = 0$. By (a) we have for $n > 0$ that

$$h(n) = \sum_{s_1+2s_2+\cdots=n} f(1)^{s_1} f(2)^{s_2}\cdots \frac{(n)_{k-1}}{s_1!s_2!\cdots}$$

$$= \sum_{k\geq 1}(n)_{k-1}[x^n]\frac{(F(x)-1)^k}{k!}$$

$$= [x^n]\int_0^{F(x)-1}(1+t)^n dt$$

$$= [x^n]\frac{F(x)^{n+1}-1}{n+1}$$

$$= [x^n]\frac{F(x)^{n+1}}{n+1}.$$

Hence by Lagrange inversion (Theorem 5.4.2, with $k=1$ and n replaced by

$n+1$) we get

$$h(n) = [x^{n+1}]\left(\frac{x}{F(x)}\right)^{\langle -1\rangle},$$

and the proof follows when char $K=0$. The case char $K=p$ is an easy consequence of the characteristic-zero case.

This result is due to R. Speicher, *Math. Ann.* **298** (1994), 611–628 (p. 616). Speicher's proof avoids the use of (a), so he in fact deduces (a) from (5.108) (see his Corollary 1).

c. This can be proved by an argument similar to (a), though the details are more complicated. The result is due to A. Nica and R. Speicher, *J. Algebraic Combinatorics* **6** (1997), 141–160 (Thm. 1.6), and is related to the "free probability theory" developed by D. V. Voiculescu. See also R. Speicher, *Mem. Amer. Math. Soc.*, vol. 132, no. 627, 1998, 88 pages, and R. Speicher, *Sém. Lotharingien de Combinatoire* (electronic) **39** (1997), B39c, 38 pp., available at http://cartan.u-strasbg.fr/~slc.

NOTE. If one defines $\zeta(n)=1$ for all n, then the function $h=f\zeta$ is as in (b). Since $\Gamma_\zeta = 1/(1+x)$, there results

$$\left(\sum_{n\geq 1} h(n)x^n\right)^{\langle -1\rangle} = \frac{1}{1+x}\left(\sum_{n\geq 1} f(n)x^n\right)^{\langle -1\rangle}.$$

It follows from the case $C(x)=1/(1+x)$ of Exercise 5.51 that this formula is equivalent to (5.108).

5.36. a. Let $u=[\frac{1}{2}(1+2x-e^x)]^{\langle -1\rangle}$ and $v=[\log(1+2x)-x]^{\langle -1\rangle}$. Thus $1+2u-2x=e^u$. If we replace u by $x+w$, then we obtain $1+2w=e^{x+w}$, whence $w^{\langle -1\rangle}=\log(1+2x)-x$. Therefore $w=v$, so $y=u-v=x$.

b. It follows from equation (5.27), equation (5.99), and part (a) of this exercise that $E_t(2x)-E_g(x)=x$, from which the proof is immediate.

c. Let us call a subset of the boolean algebra B_n of the type enumerated by $g(n)$ a *power tree*. Represent a total partition π of $[n]$ (where $n>1$) as a tree T, as in Figure 5-3. Remove any subset of the endpoints of T, in 2^n ways. The labels of the remaining vertices form a power tree. This correspondence associates each total partition of $[n]$ with 2^n power trees, such that each power tree appears exactly once, yielding (b). This elegant argument is due to C. H. Yan.

5.37. a. First note that (ii) and (iii) are obviously equivalent, since $f(u)=\log\sum_{n\geq 0} p_n(1)u^n/n!$. Given (ii), then (i) follows by expanding in powers of u both sides of the identity

$$(\exp xf(u))(\exp yf(u)) = \exp(x+y)f(u).$$

Conversely, given (i), write

$$L(x,u) = \log\sum_{n\geq 0} p_n(x)\frac{u^n}{n!}.$$

It follows from (i) that $L(x,u)+L(y,u)=L(x+y,u)$, from which it is

easy to deduce that $L(x,u) = xf(u)$ for some $f(u) = a_1u + a_2u^2 + \cdots$ (with $a_1 \neq 0$.)

For the equivalence of (i) and (iv), see G.-C. Rota and R. C. Mullin, in *Graph Theory and Its Applications* (B. Harris, ed.), Academic Press, New York, 1970, pp. 167–213 (Thm. 1) or G.-C. Rota, D. Kahaner, and A. M. Odlyzko, *J. Math. Anal. Appl.* **42** (1973), 684–760 (Thm. 1). These two papers develop a beautiful theory of "finite operator calculus" with many applications to analysis and combinatorics. For additional information and references, see S. Roman, *The Umbral Calculus*, Academic Press, Orlando, 1984. For asymptotic properties of polynomials of binomial type, see E. R. Canfield, *J. Combinatorial Theory (A)* **23** (1977), 275–290.

b.

$$\sum_n x^n \frac{u^n}{n!} = \exp xu$$

$$\sum_n (x)_n \frac{u^n}{n!} = (1+u)^x = \exp[x \log(1+u)]$$

$$\sum_n x^{(n)} \frac{u^n}{n!} = (1-u)^{-x}$$

$$= \exp[x \log(1-u)^{-1}]$$

$$\sum_n x(x-an)^{n-1} \frac{u^n}{n!} = \exp x \sum_{n\geq 1} (-an)^{n-1} \frac{u^n}{n!}$$

$$\sum_n \sum_k S(n,k) x^k \frac{u^n}{n!} = \exp x(e^u - 1)$$

$$\sum_n \sum_k \frac{n!}{k!} \binom{n+(a-1)k-1}{n-k} x^k \frac{u^n}{n!} = \exp \frac{xu}{(1-u)^a}$$

$$\sum_n \sum_k \binom{n}{k} k^{n-k} x^k \frac{u^n}{n!} = \exp xue^u.$$

A further interesting example, for which an explicit formula is not available, consists of the polynomials $n!Q_n(x)$ of Exercise 4.37. For two additional examples, see Exercise 5.38.

c. Rota and Mullin, *loc. cit.*, Thm. 2, and Rota, Kahaner, and Odlyzko, *loc. cit.*, Thm. 3.2.

d. Rota and Mullin, *loc. cit.*, Cor. 2, and Rota, Kahaner, and Odlyzko, *loc. cit.*, Cor. 3.3.

e. Let $g(u) = \sum_{n\geq 0} p_n(1)u^n/n!$, so by (a)(iii) we have $\sum_{n\geq 0} p_n(x)u^n/n! = g(u)^x$. By Exercise 5.58 there is a power series $f(u)$ satisfying

$$f(u)^x = \sum_{n\geq 0} \frac{x}{x+\alpha n} [u^n] g(u)^{x+\alpha n}$$

$$= \sum_{n\geq 0} \frac{x}{x+\alpha n} \frac{p_n(x+\alpha n)}{n!},$$

and the proof follows from (a)(iii). This result appears as part of Proposition 7.4 (p. 711) of Rota, Kahaner, and Odlyzko, *ibid.* The version of the proof given here was suggested by E. Rains.

5.38. **a.** Follows from Example 3.15.8 and condition (iii) of Exercise 5.37(a).
b. Instead of Example 3.15.8 use equation (5.77).

5.39. Let $g(n)$ (respectively, $h(n)$) be the number of series–parallel posets on $[n]$ that cannot be written as a nontrivial disjoint union (respectively, ordinal sum). Let $G(x)=\sum_{n\geq 1} g(n)x^n/n!$ and $H(x)=\sum_{n\geq 1} h(n)x^n/n!$. It is easy to see that every series–parallel poset with more than one element is either a disjoint union or ordinal sum, but not both. Hence

$$F(x) = G(x) + H(x) - x. \tag{5.133}$$

Every series–parallel poset P is a unique disjoint union $P_1+\cdots+P_k$, where each P_i is not a nontrivial disjoint union (i.e., is connected). Hence by Corollary 5.1.6,

$$1 + F(x) = e^{G(x)}. \tag{5.134}$$

Similarly P is a unique ordinal sum $P_1 \oplus \cdots \oplus P_k$, where each P_i is not a nontrivial ordinal sum. If there are exactly k summands, then by Proposition 5.1.3 the exponential generating function is $H(x)^k$. Hence

$$F(x) = \sum_{k\geq 1} H(x)^k = \frac{H(x)}{1-H(x)}. \tag{5.135}$$

It is a simple matter to eliminate $G(x)$ and $H(x)$ from (5.133), (5.134), and (5.135), thereby obtaining (5.112).

This result first appeared in R. Stanley, *Proc. Amer. Math. Soc.* **45** (1974), 295–299.

5.40. **a.** The "unlabeled" version of this problem is due to P. A. MacMahon, *The Electrician* **28** (1892), 601–602, and is further developed by J. Riordan and C. E. Shannon, *J. Math. and Physics* **21** (1942), 83–93. The labeled version given here turns out to be equivalent to the fourth problem of Schröder [**60**] discussed in the Notes. The numbers $s(n)$ satisfy $s(n)=2t(n)$ for $n \geq 2$, where $t(n)$ is the number of total partitions of an n-set, as defined in Example 5.2.5. Note also that if $f(n)$ is as in Exercise 5.26, then $f(n)=2^n s(n)$, $n \geq 1$. (See Exercise 5.36 for related results.)

The first published appearance of the formula (5.113) appears in L. Carlitz and J. Riordan, *Duke Math. J.* **23** (1955), 435–445 (eqn. (2.13)). As discussed in this reference, earlier (essentially equivalent) results were obtained by R. M. Foster (unpublished) and W. Knödel, *Monatshefte Math.* **55** (1951), 20–27. Additional aspects appear in J. Riordan, *Acta Math.* **137** (1976), 1–16. See also [**2.17**, §6.10].

b. For this result and a number of related ones, see P. J. Cameron, *Electronic J. Combinatorics* **2**, R4 (1995), 8 pp., available electronically at

http://www.combinatorics.org/Volume_2/cover.html

c. See [**3.5**, Thm. 4 and Cor. 1 on p. 351]. The table of values given in Exercise 5, p. 353, of this reference is incorrect.

5.41. **a.** Let F be a forest on the vertex set $[n]$ such that every component of F is an alternating tree rooted at some vertex i all of whose neighbors are less than i.

We obtain an alternating tree T on $\{0, 1, \dots, n\}$ by adding a vertex 0 and connecting it to the roots of the components of F. Hence if $g(n)$ denotes the number of alternating trees on the vertex set $[n]$ rooted at some vertex i all of whose neighbors are less than i, then the exponential formula (Corollary 5.1.6) yields

$$F(x) = \exp \sum_{n \geq 1} g(n) \frac{x^n}{n!}. \tag{5.136}$$

It is also easy to see that $g(n) = nf(n-1)/2$ for $n > 1$ (consider the involution on alternating trees with vertex set $[n]$ that sends vertex i to $n+1-i$), from which the stated functional equation is immediate.

Alternating trees first arose in the theory of general hypergeometric systems, as developed by I. M. Gelfand and his collaborators. In the paper I. M. Gelfand, M. I. Graev, and A. Postnikov, in *The Arnold–Gelfand Mathematical Seminars*, Birkhäuser, Boston, pp. 205–221 (§6), it is shown that $f(n)$ is the number of "admissible bases" of the space of solutions to a certain system of linear partial differential equations whose solutions are called *hypergeometric functions on the group of unipotent matrices*. The basic combinatorial properties of alternating trees were subsequently determined by A. Postnikov, *J. Combinatorial Theory (A)* **79** (1997), 360–366. See also A. Postnikov, Ph.D. thesis, Massachusetts Institute of Technology, 1997, Ch. 1.4. In particular, Postnikov established parts (a), (b), and (g) of the present exercise. Further discussion of alternating trees appears in R. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **93** (1996), 2620–2625, and A. Postnikov and R. Stanley, Deformations of Coxeter hyperplane arrangements, preprint, available at http://front.math.ucdavis.edu/math.CO/9712213

See also Exercise 6.19(p),(q).

b. Let $H(x) = x(F(x)+1)$. Then $H = x(1+e^{H/2})$, so $H(x) = [x/(1+e^{x/2})]^{\langle -1 \rangle}$. The proof follows from an application of Lagrange inversion. (See A. Postnikov, *J. Combinatorial Theory (A)* **79** (1997), 360–366, and Ph.D. thesis, Massachusetts Institute of Technology, 1997, Thm. 1.4.1, for the details.) It is an open problem to find a bijective proof.

c. This follows from equation (5.136) by reasoning as in Example 5.2.2.

d. Let $T(x) = \log F(x)^q = q \log F(x)$. It follows from (a) that

$$T(x) = \frac{qx}{2}(1 + e^{T/q}).$$

Now apply equation (5.64) to the case $F(x) = 2x/q(1 + e^{x/q})$ and $H(x) = e^x$. (Here we are using $F(x)$ and $H(x)$ in the generic sense of (5.64), and not with the specific meaning of this exercise.) This argument is due to A. Postnikov.

e. Let E be the operator on polynomials $P(q)$ defined by $EP(q) = P(q+1)$. Then (d) can be restated as

$$P_n(q) = \frac{q}{2^n}(E+1)^n q^{n-1}.$$

The proof now follows by iterating the case $\alpha = 1$ of the following lemma.

Lemma. *Let $P(q) \in \mathbb{C}[q]$ such that every zero of $P(q)$ has real part m. Let $\alpha \in \mathbb{C}$, $|\alpha| = 1$. Then every zero of the polynomial $P(q+1) + \alpha P(q)$ has real part $m - \frac{1}{2}$.*

For the history of this lemma and an elementary proof, see A. Postnikov and R. Stanley, *ibid.* (§9.3).

f. Let $R_n(q) = Q_n\left(q - \frac{n}{2}\right)$. Then $R_n(q)$ has real coefficients, is monic of degree $n-1$, and by (e) has only purely imaginary zeros (allowing 0 to be purely imaginary). Hence $R_n(q)$ has the form $q^j \prod_k \left(q^2 + a_k\right)$, $a_k \in \mathbb{R}$. Thus $R_n(-q) = (-1)^{n-1} R_n(q)$, which is equivalent to $Q_n(q) = (-1)^{n-1} Q_n(-q-n)$.

g. See A. Postnikov, *J. Combinatorial Theory (A)* **79** (1997), 360–366 (§4.1), and Ph.D. thesis, Massachusetts Institute of Technology, 1997 (§1.4.2).

h,i. The question of counting the number of regions of $\mathcal{L}_n$ was raised by N. Linial (private communication, 27 March 1995), so $\mathcal{L}_n$ is now known as the *Linial arrangement.* It was conjectured by R. Stanley that $\chi(L_n, q) = (-1)^n P_n(-q)$. This conjecture was proved by A. Postnikov, Ph.D. thesis, Massachusetts Institute of Technology, 1997 (a special case of Theorem 1.5.7), and later (using Exercise 5.50(b)) by C. A. Athanasiadis, *Advances in Math.* **122** (1996), 193–233 (Thm. 4.2). See also R. Stanley, *Proc. Nat. Acad. Sci.* **93** (1996), 2620–2625 (Cor. 4.2) and A. Postnikov and R. Stanley, *ibid.* (§9.2).

j. An alternating graph G cannot contain an odd cycle and hence is bipartite. We can partition the vertices into two sets A and B (possibly empty, and unique except for the isolated vertices of G) such that (α) every edge goes from A to B, and (β) if $i \in A$, $j \in B$, and there is an edge between i and j, then $i < j$. Call a pair (i, j) *admissible* (with respect to A and B) if $i \in A$, $j \in B$, and $i < j$. Let $h_k(n)$ be the number of ways to choose two disjoint sets A and B whose union is $\{1, 2, \ldots, n\}$, and then choose a k-element set of admissible pairs (i, j). Suppose that the elements of B are $a_1 < a_2 < \cdots < a_k$. Then the number of admissible pairs is $\nu(a_1, \ldots, a_k) = (a_1 - 1) + (a_2 - 2) + \cdots + (a_k - k)$. Hence the generating function for the subsets of such pairs according to the number of edges is $(q+1)^{\nu(a_1, \ldots, a_k)}$, so

$$\sum_k h_k(n) q^k = \sum_{1 \le a_1 < \cdots < a_k \le n} (q+1)^{\nu(a_1, \ldots, a_k)} = \sum_{0 \le b_1 \le \cdots \le b_k \le n-k} (q+1)^{b_1 + \cdots + b_k}.$$

By Proposition 1.3.19 we have that for *fixed* k,

$$\sum_{0 \le b_1 \le \cdots \le b_k \le n-k} q^{b_1 + \cdots + b_k} = \binom{n}{k}.$$

It follows that

$$\sum_k h_k(n) q^k = \sum_{k=0}^{n} \binom{n}{k}_{q+1}.$$

Now an alternating graph with r isolated vertices and k edges gets counted exactly 2^r times by $g_k(n)$ (since each isolated vertex can belong to either A or B, but there is no choice for the other vertices). Hence if $u_k(n)$ denotes the number of alternating graphs on the vertices $1, 2, \ldots, n$ with no isolated vertices and with k edges, then

$$\sum_{r=0}^{n} \binom{n}{r} 2^r \sum_k u_k(n-r) q^k = \sum_k h_k(n) q^k$$

$$\sum_{r=0}^{n} \binom{n}{r} \sum_k u_k(n-r) q^k = \sum_k g_k(n) q^k.$$

From this it is routine to deduce the stated result. The case $q = 1$ appeared in R. Stanley, Problem 10572, *Amer. Math. Monthly* **104** (1997), 168.

* **k.** Let $w_1 w_2 \cdots w_n \in \mathfrak{S}_n$. Define a tree T_w with edges labeled $1, 2, \ldots, n$ as follows: If $i < j$, then the edges labeled w_i and w_j have a common vertex if and only if the sequence $w_i w_{i+1} \cdots w_j$ is either increasing or decreasing. Then T_w is an edge-labeled alternating tree, and every such tree occurs exactly twice in this way (when $n > 1$), viz., from $w_1 w_2 \cdots w_n$ and its reverse $w_n \cdots w_2 w_1$. Hence when $n > 1$ there are $n!/2$ edge-labeled alternating trees with $n + 1$ vertices. This exercise is due to A. Postnikov (private communication, December, 1997).

5.42. a. From $y = xe^y$ we have $y' = e^y + xy'e^y$, so $xy' = xe^y/(1 - xe^y) = y/(1-y) = -1 + (1-y)^{-1}$. Thus $(1-y)^{-1} = 1 + xy' = 1 + \sum_{n \geq 1} n^n x^n / n!$.

b. Since $[1 - R(x)]^{-1} = 1 + R(x) + R(x)^2 + \cdots$, by Proposition 5.1.3 we seek a bijection $\varphi : \mathcal{R}_n^1 \cup \mathcal{R}_n^2 \cup \cdots \to \mathcal{T}_n^*$, where for $n \geq 1$ $\mathcal{R}_n^j$ is the set of j-tuples $(\tau_1, \ldots, \tau_j)$ of (nonempty) rooted trees whose total vertex set is $[n]$, and where $\mathcal{T}_n^*$ is the number of double rooted trees on $[n]$. Given $(\tau_1, \ldots, \tau_j) \in \mathcal{R}_n^j$, let v_i be the root of τ_i. Let P be a path with successive vertices $v_1, v_2, \ldots, v_j$. Label v_1 by s and v_j by e, and attach to each v_i the remainder of the tree τ_i. This yields the desired double rooted tree on $[n]$. This bijection is illustrated in Figure 5-27.

5.43. Let T be a leaf-labeled tree as in the problem. Iterate the following procedure until all vertices are labeled except the root. At the start, the leaves are labeled $1, \ldots, k$. Assume now that labels $1, 2, \ldots, m$ have been used. Label by $m + 1$

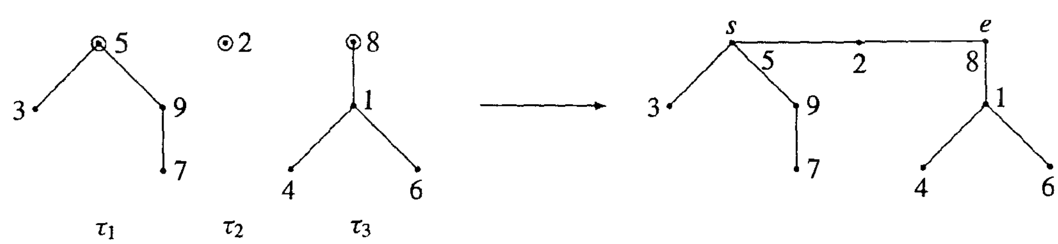


Figure 5-27. A bijection from j-tuples of rooted trees to double rooted trees.

the vertex v satisfying: (a) v is unlabeled and all successors of v are labeled, and (b) among all unlabeled vertices with all successors labeled, the vertex having the successor with the *least* label is v. Now let the blocks of the partition π consist of the labels of the successors of each nonleaf vertex v. It can be checked that this procedure yields the desired bijection.

Similar bijections appear in Erdős and Székely **[21]** and W. Y. C. Chen, *Proc. Natl. Acad. Sci. U.S.A.* **87** (1990), 9635–9639. See also W. Y. C. Chen, *Europ. J. Combinatorics* **15** (1994), 337–343. (A further bijection was discovered independently by M. Haiman.) The Erdős–Székely bijection has the minor defect of not preserving the leaf labels when the nonroot vertices are labeled. Erdős and Székely go on to deduce from their bijection many standard results on the enumeration of trees, including our Theorem 5.3.4 (or Corollary 5.3.5) and Theorem 5.3.10.

5.44. Let $r_j = \#\{i : a_i = j\}$. Given the permutation $w = w_1 \cdots w_n$, define a word $\varphi(w) = x_{m_1} \cdots x_{m_n} x_0$ as follows: If w_i is the first occurrence of a letter k, then $m_i = a_k$. Otherwise $m_i = 0$. One checks that φ is a map between the set S of permutations we wish to count and the set T of elements of the monoid $\mathcal{B}^*$ defined by equation (5.50) containing r_j copies of x_j and $1 + \sum(a_i - 1)$ copies of x_0, and that every element of T is the image of $\prod r_j!$ elements of S. The proof follows from Theorem 5.3.10. Is there a simpler proof?

An easy bijection shows that the result of this exercise is equivalent to the statement that the number of nonnesting partitions of $[n]$ (as defined in Exercise 6.19(uu)) with r_j blocks of size j is given by $n!/((n-k+1)!r_1!r_2!\cdots)$. Note the curious fact that by Exercise 5.35(a) this number is also the number of noncrossing partitions of $[n]$ with r_j blocks of size j. It is not difficult to give a bijective proof of this fact.

5.45. Let $y = \sum_{n\geq 1} t_n x^n$ and $z = \sum_{n\geq 0} f_n x^n$. It is easy to see that kxy^k is the generating function for recursively labeled trees for which the root has exactly k subtrees. Hence

$$y = x + 2xy^2 + 3xy^3 + \cdots = \frac{x}{(1-y)^2}.$$

It is then routine to use the Lagrange inversion formula to obtain the stated formula for t_n. Similarly $z = 1/(1-y)$, so $y = x/(1-y)^2 = xz^2$ and $z = 1/(1-xz^2)$. Again it is routine to use Lagrange inversion to find f_n, or to observe from $z = 1/(1-xz^2)$ that $z = 1 + xz^3$, the generating function for ternary trees. With a little more work these arguments can be "bijectivized," yielding a bijection from recursively labeled forests to ternary trees (and similarly from recursively labeled trees to pairs of ternary trees). Recursively labeled forests were first defined by A. Björner and M. L. Wachs, *J. Combinatorial Theory (A)* **52** (1989), 165–187.

5.46. Define a ternary tree $\gamma(T)$ whose vertices are the edges of T as follows. Let j be the smallest vertex of T (in this case, $j = 1$), and let k be the largest vertex for which jk is an edge e. Define three subtrees of T as follows. T_1 is the connected component containing vertex 1 of the graph $T - e$. T_2 is the connected

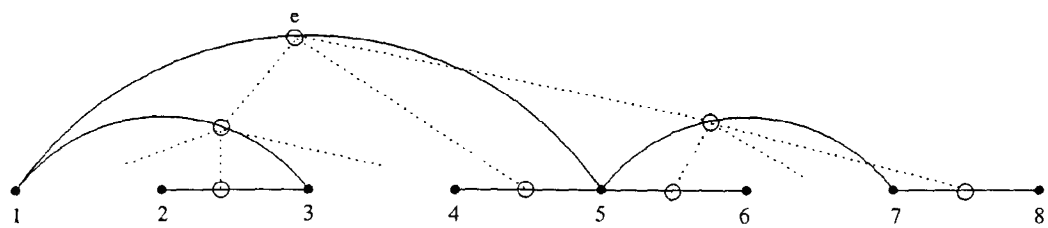


Figure 5-28. A ternary tree constructed from a noncrossing tree.

component containing vertex k of the graph obtained from T by removing edge e and vertices $k+1, k+2, \ldots, n$. T_3 is the graph obtained from T by removing vertices $1, 2, \ldots, k-1$. Define e to be the root of $\gamma(T)$, and recursively define $\gamma(T_i)$ to be the i-th subtree of the root. It is easy to see that γ is a bijection from noncrossing trees on $[n]$ to ternary trees with $n-1$ vertices. See Figure 5-28 for an example. In this figure the vertices of $\gamma(T)$ are shown as open circles and the edges as dotted lines. Three edge directions with empty subtrees have been drawn to make the ternary structure clear. Essentially the bijection just described was suggested independently by R. Simion and A. Postnikov. Noncrossing trees were first enumerated by S. Dulucq and J.-G. Penaud, *Discrete Math.* **117** (1993), 89–105 (Lemme 3.11). For further information and references, see M. Noy, *Discrete Math.* **180** (1998), 301–313. Dulucq and Penaud, *ibid.* (Proposition 2.1), also give a bijection between plane ternary trees with $n-1$ vertices and ways of drawing n chords with no common endpoints between $2n$ points on a circle such that the intersection graph G of the set of chords is a tree. (The chords are the vertices of G, with an edge connecting two vertices u and v if and only if u and v intersect (as chords).)

5.47. a. Let $w \in \mathfrak{S}_n$, and let (i, j) be a transposition in $\mathfrak{S}_n$. It is easy to see that if i and j are in different cycles of w then these two cycles are merged into a single cycle in the product $(i, j)w$. From this it follows that a product $\tau_1 \cdots \tau_{n-1}$ of $n-1$ transpositions is an n-cycle if and only if the the graph on the vertex set $[n]$ whose edges are the pairs transposed by the τ_k's is a tree. There are n^{n-2} trees on $[n]$ (Proposition 5.3.2) and $(n-1)!$ ways to linearly order their edges. Hence there are $(n-1)!n^{n-2}$ ways to write *some* n-cycle as a product of $n-1$ transpositions. By "symmetry" all $(n-1)!$ n-cycles have the same number of representations as a product of $n-1$ transpositions. Hence any particular n-cycle, such as $(1, 2, \ldots, n)$, has n^{n-2} such representations. This result is usually attributed to J. Dénes, *Publ. Math. Institute Hungar. Acad. Sci.* (= *Magyar Tud. Akad. Mat. Kutato Int. Kozl.*) **4** (1959), 63–71, and has spawned a large literature. However, a much more general theorem was announced (with a sketch of the proof) by A. Hurwitz, *Math. Ann.* **39** (1891), 1–66 (see part (c) of this exercise). Bijective proofs of this exercise were given by P. Moszkowski, *Europ. J. Combin.* **10** (1989), 13–16; I. P. Goulden and S. Pepper, *Discrete Math.* **113** (1993), 263–268; and C. M. Springer, in *Eighth International Conference on Formal Power Series and Algebraic Combinatorics, University of Minnesota, June 25–29, 1996*, pp. 427–438.

b. The formula $g(n) = \frac{1}{2n-1}\binom{3(n-1)}{n-1}$ was first proved by J. A. Eidswick, *Discrete Math.* **73** (1989), 239–243, and J. Q. Longyear, *Discrete Math.* **78** (1989), 115–118. A number of proofs were given subsequently, including I. P. Goulden and D. M. Jackson, *J. Algebra* **16** (1994), 364–378, and C. M. Springer, *ibid.* (Both these papers prove much more general results.) We sketch a bijective proof based on a suggestion of A. Postnikov. Given a noncrossing tree on $[n]$, label the edges with the labels $1, 2, \ldots, n-1$ such that the following condition holds. For every vertex i, if the vertices adjacent to i are $j_1 < \cdots < j_r < k_1 < \cdots < k_s$ with $j_r < i < k_1$, and if $\lambda(m)$ denotes the label of the edge im, then

$$\lambda(j_r) < \lambda(j_{r-1}) < \cdots < \lambda(j_1) < \lambda(k_s) < \lambda(k_{s-1}) < \cdots < \lambda(k_1).$$

Let τ_i be the transposition (a, b), where ab is the edge of T labeled i. Then it is not hard to show that $\tau_1\tau_2\cdots\tau_{n-1} = (1, 2, \ldots, n-1)$ and that each equivalence class is obtained exactly once in this way, thus giving the desired bijection.

c. This result was stated with a sketch of a proof by A. Hurwitz in 1891 (reference in (a)). The first complete proof was given by I. P. Goulden and D. M. Jackson, *Proc. Amer. Math. Soc.* **125** (1997), 51–60, based on the theory of symmetric functions. A reconstruction of the proof of Hurwitz, together with much interesting further information, was given by V. Strehl, *Sém. Lotharingien de Combinatoire* (electronic) **37** (1996), B37c, 12 pp., available at http://cartan.u-strasbg.fr/~slc. A direct combinatorial proof would be highly desirable. Some further aspects of "transitive factorizations" are discussed in I. P. Goulden and D. M. Jackson, Transitive factorisations in the symmetric group, and combinatorial aspects of singularity theory, Research Report 97-13, Department of Combinatorics and Optimization, University of Waterloo, July 1997.

5.48. a. Let G be a connected graph on $[n]$. Define a certain spanning tree τ_G of G as follows. Start at vertex 1, and always move to the *greatest* adjacent unvisited vertex if there is one; otherwise backtrack. Stop when every vertex has been visited, and let τ_G consist of the vertices and edges visited. We leave to the reader the proof of the following crucial lemma.

Lemma. *Let τ be a tree on $[n]$. A connected graph G satisfies $\tau_G = \tau$ if and only if τ is a spanning tree of G, and every other edge of G has the form $\{i, k\}$, where (i, j) is an inversion of τ and k is the unique predecessor of j in the rooted tree (with root 1) τ.*

Thus the $n-1$ edges of τ must be edges of G, while any subset of the $i(\tau)$ "inversion edges" defined by the previous lemma may constitute the remaining edges of G. Hence

$$\sum_G t^{e(G)} = t^{n-1}(1+t)^{\mathrm{inv}(\tau)}, \tag{5.137}$$

where G ranges over all connected graphs on $[n]$ satisfying $\tau_G = \tau$. Summing (5.137) over all τ completes the proof.

Equation (5.115) was first proved using an indirect generating function method by C. L. Mallows and J. Riordan, *Bull. Amer. Math. Soc.* **74** (1968), 92–94. (See also [**47**, §4.5].) The elegant proof given here is due to I. M. Gessel and D.-L. Wang, *J. Combinatorial Theory (A)* **26** (1979), 308–313. Gessel and Wang also give a similar result related to the enumeration of acyclic digraphs and tournaments. For some further results related to inversions in trees, see G. Kreweras, *Period. Math. Hungar.* **11**(4) (1980), 309–320, and J. S. Beissinger, *J. Combinatorial Theory (B)* **33** (1982), 87–92, as well as Exercises 5.49(c) and 5.50(d). A remarkable conjectured connection between tree inversions and invariant theory appears in M. Haiman, *J. Algebraic Combinatorics* **3** (1994), 17–76 (§2.3).

b. Substitute $t-1$ for t in equation (5.115), take the logarithm of both sides, and differentiate with respect to x. An explicit statement of the formula appears in [**28**, (14.7)]. For a generalization, see R. Stanley, in *Mathematical Essays in Honor of Gian-Carlo Rota* (B. E. Sagan and R. P. Stanley, eds.), Birkhäuser, Boston/Basel/Berlin, 1998, pp. 359–375, Thm. 3.3.

5.49. a. If some $b_i > i$, then at least $n-i+1$ cars prefer the $n-i$ spaces $i+1, i+2, \ldots, n$ and hence are unable to park. Thus the stated condition is necessary. The sufficiency can be proved by induction on n. Namely, suppose that $a_1 = k$. Define for $1 \le i \le n-1$,

$$a'_i = \begin{cases} a_{i+1} & \text{if } a_{i+1} \le k \\ a_{i+1} - 1 & \text{if } a_{i+1} > k. \end{cases}$$

Then the sequence $(a'_1, \ldots, a'_{n-1})$ satisfies the condition so by induction is a parking function. But this means that for the original sequence $\alpha = (a_1, \ldots, a_n)$, the cars $C_2, \ldots, C_n$ can park after car C_1 occupies space k. Hence α is a parking function, and the proof follows by induction (the base case $n=1$ being trivial).

b. Add an additional parking space 0 after space n, and allow 0 also to be a preferred parking space. Consider the situation where the cars $C_1, \ldots, C_n$ enter the street as before (beginning with space 1), but if a car is unable to park it may start over again at 1 and take the first available space. Of course now every car can park, and there will be exactly one empty space. If the preferences $(a_1, \ldots, a_n)$ lead to the empty space i, then the preferences $(a_1+k, \ldots, a_n+k)$ will lead to the empty space $i+k$ (addition in G). Moreover, α is a parking function if and only if the space 0 is left empty. From this the proof follows.

Parking functions were first considered by A. G. Konheim and B. Weiss, *SIAM J. Applied Math.* **14** (1966), 1266–1274, in connection with a hashing problem. They proved the formula $P(n) = (n+1)^{n-1}$ using recurrence relations. (The characterization (a) of parking functions seems to be part of the folklore of the subject.) The elegant proof given here is due to H. Pollak, described in J. Riordan, *J. Combinatorial Theory* **6** (1969), 408–411, and D. Foata and J. Riordan, *Aequationes Math.* **10** (1974), 10–22 (p. 13). Some bijections between parking functions and trees on the vertex set $[n+1]$ appear in the previous reference, as well as in J. Françon, *J. Combinatorial Theory (A)* **18** (1975), 27–35; P. Moszkowski, *Period. Math. Hungar.* **20** (1989), 147–154 (§3); and J. S. Beissinger and U. N. Peled, *Electronic J. Combinatorics*

4(2), R4 (1997), 10 pp., available electronically at

http://www.combinatorics.org/Volume_4/wilftoc.html

c. This result is due to G. Kreweras, *Period. Math. Hungar.* **11**(4) (1980), 309–320. Kreweras deals with *suites majeures* (major sequences), which are obtained from parking functions $(a_1, \ldots, a_n)$ by replacing a_i with $n + 1 - a_i$.

d. Suppose that cars $C_1, \ldots, C_{i-1}$ have already parked at spaces $u_1, \ldots, u_{i-1}$. Then C_i parks at u_i if and only if spaces $a_i, a_i+1, \ldots, u_i - 1$ are already occupied. Thus a_i can be any of the numbers $u_i, u_i - 1, \ldots, u_i - \tau(u, u_i) + 1$. There are therefore $\tau(u, u_i)$ choices for a_i, so

$$\nu(u) = \tau(u, u_1) \cdots \tau(u, u_n) = \tau(u, 1) \cdots \tau(u, n).$$

This result is implicit in Konheim and Weiss, *ibid.*

e. Given σ, define a poset $(P_\sigma, \stackrel{\sigma}{<})$ on $[n]$ by the condition that $j \stackrel{\sigma}{<} i$ if either $0 < i - j \leq s_i$ or $0 < j - i \leq t_i$. It is easy to see that P_σ is a tree, and that T_σ consists of the linear extensions of P_σ (where we regard a linear extension of P_σ as a permutation of its elements). By definition of P_σ we have $\#\Lambda_i = s_i + t_i$, where $\Lambda_i = \{j \in P_\sigma : j \stackrel{\sigma}{\leq} i\}$, and the proof follows from Supplementary Problem 3.1(b).

NOTE. The trees T_σ by definition have the property that the elements of Λ_i form a set of consecutive integers. Hence T_σ is a *recursively labeled tree* in the sense of Exercise 5.45. There follows from Theorem 2.2 of the paper of Björner and Wachs cited there the curious result

$$\sum_{u \in T_\sigma} q^{\mathrm{inv}(u)} = \sum_{u \in T_\sigma} q^{\mathrm{maj}(u)}.$$

Note that this formula is a refinement of the fact that maj and inv are equidistributed over $\mathfrak{S}_n$ (Corollary 1.3.10 and Corollary 4.5.9).

f. **First solution.** Suppose that $\alpha_1, \ldots, \alpha_k$ is a sequence of prime parking functions, where the length of α_i is d_i. Let β_i denote α_i with $d_1 + d_2 + \cdots + d_{i-1}$ added to each term. Then any permutation of all the terms of all the β_i's is a parking function, and conversely given any parking function one can uniquely reconstruct $\alpha_1, \ldots, \alpha_k$. From this it follows (using equation (5.116)) that

$$\sum_{n \geq 0} (n+1)^{n-1} \frac{x^n}{n!} = \frac{1}{1 - \sum_{n \geq 1} Q(n) \frac{x^n}{n!}}.$$

The proof now follows from equation (5.67). The definition of prime parking functions and the above proof of their enumeration is due to I. Gessel (private communication, 1997).

Second solution. Let r_i be the number of entries of α equal to i. One checks that the parking function α is prime if and only if every partial sum of the sequence $r_1 - 1, r_2 - 1, \ldots, r_{n-1} - 1$ is positive (in which case $r_n = 0$ and $\sum_i (r_i - 1) = 1$). A version of Lemma 5.3.7 shows that any sequence of integers with sum 1 has exactly one cyclic permutation all of whose partial sums are positive. From this it follows that if we regard the elements of the group $L = \mathbb{Z}/(n-1)\mathbb{Z}$ as being the integers $1, 2, \ldots, n-1$, then every coset

of the subgroup M of L^n generated by $(1, 1, \ldots, 1)$ contains exactly one prime parking function. Hence $Q(n) = [L^n : M] = (n-1)^{n-1}$. This argument is due to L. Kalikow.

5.50. a. The number of regions was first computed by J.-Y. Shi, *Lecture Notes in Mathematics* **1179**, Springer, Berlin/Heidelberg/New York, 1986, Ch. 7, and *J. London Math. Soc.* **35** (1987), 56–74. For this reason the arrangement $\mathcal{S}_n$ is called the *Shi arrangement*. A more elementary (though nonbijective) proof was subsequently given by P. Headley, Ph.D. thesis, University of Michigan, Ann Arbor, 1994, Ch. VI, and *Formal Power Series and Algebraic Combinatorics, FPSAC '94, May 23–27, 1994*, DIMACS preprint, pp. 225–232 (§5). For a simple nonbijective proof, see the solution to (c). A bijection between the regions of $\mathcal{S}_n$ and parking functions of length n (as defined in Exercise 5.49) is due to I. Pak and R. Stanley, stated in R. Stanley, *Proc. Nat. Acad. Sci.*, **93** (1996), 2620–2625 (Thm. 5.1) and proved in R. Stanley, in *Mathematical Essays in Honor of Gian-Carlo Rota* (B. E. Sagan and R. P. Stanley, eds.), Birkhäuser, Boston/Basel/Berlin, 1998, pp. 359–375, Thm. 2.1. This bijection has the virtue of allowing an easy proof of (c). Simpler bijections lacking this property were given by J. Lewis, Parking functions and regions of the Shi arrangement, preprint dated August 1, 1996, and C. A. Athanasiadis and S. Linusson, *Discrete Math.*, to appear.

b. Let L_p denote the intersection poset of the arrangement $\mathcal{A}_p$. The poset L_p is determined by the vanishing of certain minors of the coefficient matrix of the hyperplanes in $\mathcal{A}$. Hence $L_p \cong L$ for $p \gg 0$. Let $\bar{L}_p$ denote L_p with a $\hat{1}$ adjoined. For $V \in \bar{L}_p$, let $f(V)$ be the number of points $v \in \mathbb{F}_p^n$ such that V is the largest element (i.e., the least element under inclusion, since L_p is ordered by *reverse* inclusion) of $\bar{L}_p$ for which $v \in V$. In particular, $f(\hat{1}) = 0$. Clearly

$$\#V = p^{\dim V} = \sum_{W \geq V \text{ in } \bar{L}_p} f(W).$$

Möbius inversion (Proposition 3.7.2) yields

$$f(V) = \sum_{W \geq V} \mu(V, W) p^{\dim W},$$

where μ denotes the Möbius function of $\bar{L}$. Setting $V = \hat{0}$ completes the proof.

This result is implicit in H. Crapo and G.-C. Rota, *On the Foundations of Combinatorial Theory: Combinatorial Geometries*, preliminary edition, MIT Press, Cambridge, Massachusetts, 1970 (see pp. 193–194 of C. A. Athanasiadis, *Advances in Math.* **122** (1996), 193–233, for an explanation), and P. Orlik and H. Terao, *Arrangements of Hyperplanes*, Springer-Verlag, Berlin, 1992, Thm. 2.3.22. It was first stated explicitly by C. A. Athanasiadis, Ph.D. thesis, Massachusetts Institute of Technology, 1996, Thm. 5.2.1. Athanasiadis was the first person to use this result systematically to compute characteristic polynomials. See his papers *Advances in Math.* **122** (1996), 193–233, and MSRI Preprint 1997-059, Mathematical Sciences Research Institute, Berkeley, CA.

c. We want to compute the number of n-tuples $(x_1, \ldots, x_n) \in \mathbb{F}_p^n$ such that if $i < j$, then $x_i \neq x_j$ and $x_i \neq x_j + 1$. There are $(p-n)^{n-1}$ ways to choose a weak ordered partition $\pi = (B_1, \ldots, B_{p-n})$ of $[n]$ into $p-n$ blocks such that $1 \in B_1$. Choose x_1 in p ways. Think of the elements of $\mathbb{F}_p$ as being arranged in a circle, in the clockwise order $0, 1, \ldots, p-1$. We will place the numbers $1, 2, \ldots, n$ on some of the p points of this circular depiction of $\mathbb{F}_p$. Place the elements of B_1 consecutively in increasing order when read clockwise, with 1 placed at x_1. Then skip one space (in clockwise order) and place the elements of B_2 consecutively in increasing order. Then skip one space and place the elements of B_3 consecutively in increasing order, etc. Let x_i be the point at which i is placed. It is easy to see that this gives a bijection between the $p(p-n)^{n-1}$ choices of (π, x_1) and the allowed values of $(x_1, \ldots, x_n)$, so the proof follows from (b). This argument is due to C. A. Athanasiadis, Ph.D. thesis, Massachusetts Institute of Technology, 1996, Thm. 6.2.1, and *Advances in Math.* **122** (1996), 193–233, Thm. 3.3. The characteristic polynomial of the Shi arrangement was first computed by P. Headley, Ph.D. thesis, University of Michigan, Ann Arbor, 1994, Ch. VI; *Formal Power Series and Algebraic Combinatorics, FPSAC '94, May 23–27, 1994*, DIMACS preprint, pp. 225–232 (§5); and *J. Algebraic Combinatorics* **6** (1997), 331–338 (Thm. 2.4 in the case $\Phi = A_n$), by a different method. A further proof appears in A. Postnikov, Ph.D. thesis, Massachusetts Institute of Technology., 1997 (Example 1, p. 39), and A. Postnikov and R. Stanley, Deformations of Coxeter hyperplane arrangements, preprint available at http://front.math.ucdavis.edu/math.CO/9712213 (Cor. 9.3).

*

d. This result is equivalent to a theorem of I. Pak and R. Stanley that is stated in R. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **93** (1996), 2620–2625 (Thm. 5.1), and proved in R. Stanley, in *Mathematical Essays in Honor of Gian-Carlo Rota* (B. Sagan and R. Stanley, eds.), Birkhäuser, Boston/Basel/Berlin, 1998, pp. 359–375 (the case $k=1$ of Cor. 2.20).

e. Let $x = (x_1, \ldots, x_n)$ belong to some region R of $\mathcal{S}_n$. Define $\pi_x \in \mathfrak{S}_n$ by the condition

$$x_{\pi_x(1)} > x_{\pi_x(2)} > \cdots > x_{\pi_x(n)}.$$

Let $I_x = \{(i,j) : 1 \leq i < j \leq n,\ x_j + 1 > x_i > x_j\}$. It is not difficult to show that the map $R \mapsto (\pi_x, I_x)$ is a bijection between the regions of $\mathcal{S}_n$ and the pairs (π_x, I_x) where $\pi_x \in \mathfrak{S}_n$ and $I_x \in J(P_{\pi_x})$. Moreover, $d(R) = \binom{n}{2} - |I_x|$, and the proof follows. This result was stated without proof in R. Stanley, *ibid.* (after Theorem 5.1).

f. Let $\pi = \{B_1, \ldots, B_{n-k}\}$ be a partition of $[n]$, and let w_i be a permutation of B_i. Let X consist of all points $(x_1, \ldots, x_n)$ in $\mathbb{R}^n$ such that $x_a - x_b = m$ if a and b appear in the same block B_i of π, a appears to the left of b in w_i, and there are exactly m ascents appearing in w_i between a and b. For instance, if $w_1 = 495361$ and $w_2 = 728$, then X is defined by the conditions

$$x_4 = x_9 + 1 = x_5 + 1 = x_3 + 1 = x_6 + 2 = x_1 + 2, \qquad x_7 = x_2 = x_8 + 1.$$

This defines a bijection between the partitions of $[n]$ into $n-k$ linearly ordered blocks and the elements X of $L_{\mathcal{S}_n}$ of rank k.

5.51. Assume (i). Substituting $A(x)$ for x yields

$$\frac{x}{C(A(x))} = B^{\langle -1\rangle}(A(x)).$$

Substituting $B(x)$ for x in (i) yields

$$A^{\langle -1\rangle}(B(x)) = xC(B(x)).$$

But $\left[A^{\langle -1\rangle}(B(x))\right]^{\langle -1\rangle} = B^{\langle -1\rangle}(A(x))$, so (ii) follows. The steps are reversible, so (i) and (ii) are equivalent.

5.52. **a.** See [**2.3**, Ch. 3.7] where also the polynomials $\varphi_n(k)$ are given for $n \le 7$.

b. First check that for fixed n, the quantities $[x^n]F^{\langle j+k\rangle}(x)$ and $[x^n]F^{\langle j\rangle}(F^{\langle k\rangle}(x))$ are polynomials in j and k. Since these two polynomials agree for all $j, k \in \mathbb{P}$, they must be the same polynomials. A similar argument works for the second identity. See [**2.3**, Thm. B, p. 148].

* **5.53.** We need to compute

$$f(n) := [x^{n-1}]\left(1 - \tfrac{1}{2}x\right)^{-n}(1-x)^{-1} \qquad \text{[why?]}.$$

In equation (5.64), let $x/F(x) = (1 - \frac{1}{2}x)^{-1}$ and $H'(x) = (1-x)^{-1}$. Then

$$F^{\langle -1\rangle}(x) = 1 - \sqrt{1-2x}, \qquad H(x) = -\log(1-x),$$

so

$$\begin{aligned} f(n) &= n[x^n](-\log\sqrt{1-2x}) \\ &= -\frac{n}{2}[x^n]\log(1-2x) \\ &= 2^{n-1}, \end{aligned}$$

exactly half the sum of the entire series.

This result is equivalent to the identity

$$2^{n-1} = \sum_{j=0}^{n-1} 2^{-j}\binom{n+j-1}{j},$$

or equivalently (putting $n+1$ for n)

$$4^n = \sum_{j=0}^{n} 2^{n-j}\binom{n+j}{j}.$$

Is there a simple combinatorial proof?

Bromwich [**5**, Example 20, p. 199] attributes the result of this exercise to *Math. Trip.* 1903.

5.54. By equation (5.53) we have

$$[x^{-1}]F(x)^{-n} = n[x^n]F^{\langle -1\rangle}(x).$$

The compositional inverses of the four functions are given by

$$\sin^{-1} x = \sum_{m\geq 0} 4^{-m}\binom{2m}{m}\frac{x^{2m+1}}{2m+1}$$

$$\tan^{-1} x = \sum_{m\geq 0}(-1)^m \frac{x^{2m+1}}{2m+1}$$

$$e^x - 1 = \sum_{n\geq 1}\frac{x^n}{n!}$$

$$x + \frac{x^2}{2(1-x)} = x + \frac{1}{2}\sum_{n\geq 2} x^n,$$

yielding the four answers

$$\begin{cases} 0, & n = 2m \\ 4^{-m}\binom{2m}{m}, & n = 2m+1 \end{cases}$$

$$\begin{cases} 0, & n = 2m \\ (-1)^m, & n = 2m+1 \end{cases}$$

$$\frac{1}{(n-1)!}$$

$$\begin{cases} 1, & n = 1 \\ n/2, & n \geq 2. \end{cases}$$

Bromwich [5, Example 19, p. 199] attributes these formulas to Wolstenholme.

5.55. a. Let $y \in \mathbb{Q}[[x]]$ satisfy $y = xF_1(y)$. By (5.55) with $k = 1$ we have

$$n[x^n]y = [x^{n-1}]F_1(y)^n = 1,$$

so $y = \sum_{n\geq 1} x^n/n = -\log(1-x)$. Hence $y^{\langle -1\rangle} = 1 - e^{-x}$, so $F_1(x) = x/y^{\langle -1\rangle} = x/(1-e^{-x})$.

Note. The *Bernoulli numbers* B_n are defined by

$$\frac{x}{e^x - 1} = \sum_{n\geq 0} B_n \frac{x^n}{n!}.$$

Hence

$$F_1(x) = \sum_{n\geq 0} (-1)^n B_n \frac{x^n}{n!}.$$

Essentially the same result is attributed to Wolstenholme and *Math. Trip.* 1904

by Bromwich [5, Example 18, p. 199]. Somewhat surprisingly, this result has applications to algebraic geometry. See Lemma 1.7.1 of F. Hirzebruch, *Topological Methods in Algebraic Geometry*, Springer-Verlag, New York, 1966.

A more general result is the following: Given $f(x)=\sum_{n\geq 1}\alpha_{n-1}(x^n/n)\in\mathbb{C}[[x]]$, it follows from (5.57) that the unique power series $F(x)$ satisfying $[x^n]F(x)^{n+1}=\alpha_n$ for all $n\in\mathbb{N}$ is given by $F(x)=x/f^{\langle -1\rangle}(x)$.

b,c. Note that $F_k(0)=1$ (the case $n=0$). Let $G_k(x)=x/F_k(x^k)$. The condition on $F_k(x)$ becomes

$$[x^n]\left(\frac{x}{G_k(x)}\right)^{n+1}=\begin{cases}1 & \text{if } n\equiv 0 \pmod{k}\\ 0 & \text{otherwise.}\end{cases}$$

By Lagrange inversion (Theorem 5.4.2) we have

$$[x^n]\left(\frac{x}{G_k(x)}\right)^{n+1}=(n+1)[x^{n+1}]G_k^{\langle -1\rangle}(x).$$

Hence

$$G_k^{\langle -1\rangle}(x)=\sum_{m\geq 0}\frac{x^{km+1}}{km+1}.$$

When $k=2$ we have $G_2^{\langle -1\rangle}=\frac{1}{2}\log\frac{1+x}{1-x}$, whence $G_2(x)=(e^{2x}-1)/(e^{2x}+1)$ and $F_2(x)=\sqrt{x}/(\tanh\sqrt{x})$. This result appears as Lemma 1.5.1 of Hirzebruch, *ibid*.

When $k>2$ there is no longer a simple way to invert the series $G_k^{\langle -1\rangle}(x)=\sum_{m\geq 0}x^{km+1}/(km+1)$.

5.56. a. First Solution. More generally, let $G(x)=\sum_{n\geq 0}b_nx^n$ be any power series with $b_0=1$. Define

$$H(x)=x\exp\sum_{n\geq 1}b_n\frac{x^n}{n},$$

so

$$\frac{H'(x)}{H(x)}=\frac{G(x)}{x}.$$

Set $y=F(x)=a_1x+a_2x^2+\cdots$, $a_1\neq 0$. Consider the formal power series

$$\log\frac{H(y^{\langle -1\rangle})}{x}:=\sum_{i\geq 1}p_ix^i.$$

Then

$$\log \frac{H(x)}{y} = \sum_{i\geq 1} p_i y^i$$

$$\Longrightarrow \qquad \frac{H'(x)}{H(x)} - \frac{y'}{y} = \sum_{i\geq 1} i p_i y^{i-1} y'$$

$$\Longrightarrow y^{-n}\left(\frac{G(x)}{x} - \frac{y'}{y}\right) = \sum_{i\geq 1} i p_i y^{i-n-1} y'. \tag{5.138}$$

Take the coefficient of $1/x$ on both sides. As in the first proof of Theorem 5.4.2, we obtain

$$[x^{-1}]\frac{G(x)}{xy^n} = np_n. \tag{5.139}$$

Now take $G(x) = 1$ in (5.139), so $H(x) = x$. We get

$$n[x^n]\log \frac{y^{\langle -1\rangle}}{x} = np_n = [x^{-1}]\frac{1}{xy^n} = [x^n]\left(\frac{x}{y}\right)^n,$$

as desired.

Second Solution. Define $H(x) = \log[x/F(x)]$. Then (5.64) becomes

$$\begin{aligned} n[x^n]\log \frac{F^{\langle -1\rangle}(x)}{x} &= [x^{n-1}]\left(\frac{1}{x} - \frac{F'(x)}{F(x)}\right)\left(\frac{x}{F(x)}\right)^n \\ &= [x^n]\left(\frac{x}{F(x)}\right)^n - [x^{-1}]\frac{F'(x)}{F(x)^{n+1}} \\ &= [x^n]\left(\frac{x}{F(x)}\right)^n + \frac{1}{n}[x^{-1}]\frac{d}{dx}F(x)^{-n} \\ &= [x^n]\left(\frac{x}{F(x)}\right)^n, \end{aligned}$$

as desired.

Third Solution. Equation (5.53) can be rewritten

$$n[x^n]\frac{1}{k}\left(\frac{F^{\langle -1\rangle}(x)}{x}\right)^k = [x^n]\left(\frac{x}{F(x)}\right)^n. \tag{5.140}$$

The first proof of Theorem 5.4.2 is actually valid for any $k \in \mathbb{R}$, so we can let $k \to 0$ in (5.140) to get (after some justification) equation (5.118).

The result of this exercise goes back to J. L. Lagrange, *Mém. Acad. Roy. Sci. Belles-Lettres Berlin* **24** (1770); *Oeuvres*, Vol. 3 Gauthier-Villars, Paris, 1869, pp. 3–73. It was rediscovered by I. Schur, *Amer. J. Math.* **69** (1947), 14–26.

b. Let $G(x)=x/F(x)$. By (a),

$$\delta_{0n} = [x^n]G(x)^n = n[x^n]\log\frac{F^{\langle -1\rangle}(x)}{x}.$$

Thus $x = \log[F^{\langle -1\rangle}(x)/x]$, so $F^{\langle -1\rangle}(x)=xe^x$. Hence

$$\begin{aligned} G(x) &= x/(xe^x)^{\langle -1\rangle} \\ &= 1+\sum_{n\geq 1}(-1)^{n-1}(n-1)^{n-1}\frac{x^n}{n!} \qquad (\text{with } 0^0=1), \end{aligned}$$

by a simple application of (5.53) (or by substituting $-x$ for x in (5.67)).

5.57. In Corollary 5.4.3 take $H(x)=\log(1+x)$ and $x/F(x)=(1+x)^2/(2+x)$. Then $F(x)=1-(1+x)^{-2}$, so $F^{\langle -1\rangle}(x)=(1-x)^{-1/2}-1$. Equation (5.64) becomes

$$n[x^n]\tfrac{1}{2}\log(1-x)^{-1} = [x^{n-1}](1+x)^{2n-1}(2+x)^{-n}.$$

But $[x^n]\log(1-x)^{-1}=1/n$, and the result follows.

By expanding $(1+x)^{2n-1}$ and $(2+x)^{-n}$ and taking the coefficient of x^{n-1} in their product, we see that an equivalent result is the identity (replacing n by $n+1$)

$$4^n = \sum_{j=0}^{n}(-1)^{n-j}2^j\binom{2n+1}{j}\binom{2n-j}{n}.$$

Bromwich [5, Example 18, p. 199] attributes this result to *Math. Trip.* 1906.

5.58. Let $F(x)=xf(x)^\alpha$ and $G(x)=g(x)^\alpha$. Then the functional equation (5.119) becomes $F(x)=xG(F(x))$, so by ordinary Lagrange inversion (Theorem 5.4.2) we get

$$m[x^m]F(x)^k = k[x^{m-k}]G(x)^m$$

for any nonnegative integers m and k. In terms of f and g this is

$$m[x^{m-k}]f(x)^{\alpha k} = k[x^{m-k}]g(x)^{\alpha m}.$$

Now set $k=t/\alpha$ and $m=\frac{t}{\alpha}+n$, so that $t=\alpha k$ and $n=m-k$. We get the desired formula with the restriction that t/α must be a nonnegative integer. However, since both sides are polynomials in t, the formula holds for all t.

This result is due to E. Rains (private commmunication), and the above proof was provided by I. Gessel. For an application, see Exercise 5.37(e).

5.59. Define $g(x,y)\in K[[x,y]]$ to be the (unique) power series satisfying the functional equation $g=yF(x,g)$. Thus $g(x,1)=f(x)$. By Lagrange inversion (Theorem 5.4.2) we have $n[y^n]g(x,y)^k=k[y^{n-k}]F(x,y)^n$. Hence

$$g(x,t)^k = \sum_{n\geq 1}\left([y^n]g(x,y)^k\right)t^n = \sum_{n\geq 1}\frac{k}{n}\left([u^{n-k}]F(x,u)^n\right)t^n.$$

Setting $t=1$ yields the desired result. This argument is due to I. Gessel.

5.60. **a.** One method of proof is to let $B(x)=A(x)-1$ and write

$$A(x)^n = [1+B(x)]^n = \sum_{j\geq 0}\binom{n}{j}B(x)^j.$$

Thus (since $\deg B(x)^j \geq j$),

$$[x^k]A(x)^n = \sum_{j=0}^{k} \binom{n}{j} [x^k]B(x)^j,$$

which is clearly a polynomial in n of degree $\leq k$.

Alternatively, let Δ be the difference operator with respect to the variable n. Then by equation (1.26) we have

$$\begin{aligned}\Delta^{k+1}[x^k]A(x)^n &= [x^k]\sum_{i=0}^{k+1}(-1)^{k+1-i}\binom{k+1}{i}A(x)^{n+i}\\ &= [x^k]A(x)^n(A(x)-1)^{k+1}\\ &= 0.\end{aligned}$$

Now use Proposition 1.4.2(a).

b. Since $e^{tF(x)} = \sum_{n\geq 0} t^n F(x)^n/n!$, we have

$$\begin{aligned}p_k(n) &= \left[t^n \frac{x^{n+k}}{(n+k)!}\right] e^{tF(x)}\\ &= \frac{(n+k)!}{n!}\,[x^{n+k}]F(x)^n\\ &= (n+k)_k\,[x^k]\left(\frac{F(x)}{x}\right)^n. \qquad (5.141)\end{aligned}$$

Now use (a).

c. We have, as in (b),

$$\begin{aligned}\left[t^n \frac{x^{n+k}}{(n+k)!}\right] e^{tF^{\langle -1\rangle}(x)} &= (n+k)_k[x^{n+k}]F^{\langle -1\rangle}(x)^n\\ &= (n+k)_k \frac{n}{n+k}[x^k]\left(\frac{F(x)}{x}\right)^{-n-k} \qquad \text{(by (5.53))}\\ &= \frac{(n+k-1)_k}{(-n)_k}\,p_k(-n-k) \qquad \text{(by (5.141))}\\ &= (-1)^k p_k(-n-k),\end{aligned}$$

as desired.

d. *Answer:* We have $p_k(n) = S(n+k,n)$ and $(-1)^k p_k(-n-k) = s(n+k,n)$. The *Stirling number reciprocity* $S(-n,-n-k) = (-1)^k s(n+k,n)$ is further discussed in I. Gessel and R. Stanley, *J. Combinatorial Theory (A)* **24** (1978), 24–33, and D. E. Knuth, *Amer. Math. Monthly* **99** (1992), 403–422.

e. It follows from Exercise 5.17(b) that

$$p_k(n) = \frac{(n+k)!}{n!}\binom{n+k-1}{n-1}$$
$$= \frac{(n+k)(n+k-1)^2(n+k-2)^2\cdots(n+1)^2n}{(k-1)!} \qquad (k \geq 1).$$

Since $F^{\langle -1\rangle}(x) = x/(1+x) = -F(-x)$, it follows from (c) that $p_k(-n-k) = p_k(n)$. For further information on power series $F(x)$ satisfying $F^{\langle -1\rangle}(x) = -F(-x)$, see Exercise 1.41.

f. *Answer:*

$$p_k(n) = (-1)^k\binom{n+k}{k}n^k.$$

Thus $(-1)^k p_k(-n-k) = \binom{n+k-1}{k}(n+k)^k$, so

$$\exp t(xe^{-x})^{\langle -1\rangle} = \sum_{n\geq 0}\sum_{k\geq 0}\binom{n+k-1}{k}(n+k)^k t^n \frac{x^{n+k}}{(n+k)!}.$$

This formula is also immediate from Propositions 5.3.1 and 5.3.2.

5.61. a. Clearly $\mu(\bar{P}\times\bar{Q}) = \mu(\bar{P})\mu(\bar{Q})$ by Proposition 3.8.2. Now we have the disjoint union

$$\bar{P}\times\bar{Q} = (P\times Q)\,\dot\cup\,\{(x,\hat{0}_{\bar{Q}}) : x\in P\}\,\dot\cup\,\{(\hat{0}_{\bar{P}},y) : y\in Q\}\,\dot\cup\,\{(\hat{0}_{\bar{P}},\hat{0}_{\bar{Q}})\}.$$

Write $[u,v]_R$ for the interval $[u,v]$ of the poset R. If $t\in P\times Q$, then the intervals $[t,\hat{1}]_{P\times Q}$ and $[t,\hat{1}]_{\bar{P}\times\bar{Q}}$ are isomorphic. If $x\in P$, then the interval $[(x,\hat{0}_{\bar{Q}}),\hat{1}_{\bar{P}\times\bar{Q}}]$ is isomorphic to $[x,\hat{1}]_P\times\bar{Q}$. Similarly if $y\in Q$, then $[(\hat{0}_{\bar{P}},y),\hat{1}]_{\bar{P}\times\bar{Q}} \cong \bar{P}\times[y,\hat{1}]_Q$. Hence

$$\begin{aligned}
0 &= \sum_{t\in\bar{P}\times\bar{Q}} \mu_{\bar{P}\times\bar{Q}}(t,\hat{1})\\
&= \sum_{t\in P\times Q}\mu_{P\times Q}(t,\hat{1}) + \left(\sum_{x\in P}\mu_P(x,\hat{1})\right)\mu(\bar{Q})\\
&\quad + \mu(\bar{P})\left(\sum_{y\in Q}\mu_Q(y,\hat{1})\right) + \mu(\bar{P})\mu(\bar{Q})\\
&= -\mu(\overline{P\times Q}) - \mu(\bar{P})\mu(\bar{Q}) - \mu(\bar{P})\mu(\bar{Q}) + \mu(\bar{P})\mu(\bar{Q}),
\end{aligned}$$

and the proof follows. Note that Exercise 3.69(d) is a special case.

b. In Corollary 5.5.5 put $f(i) = -\mu_i = -\mu(\bar{Q}_i)$. If type $\pi = (a_1,\ldots,a_n)$ then by property (E3) of exponential structures and (a), we have $f(1)^{a_1}\cdots f(n)^{a_n} = -\mu(\overline{Q_1^{a_1}\times\cdots\times Q_n^{a_n}}) = -\mu(\hat{0},\pi)$. Hence

$$h(n) = -\sum_{\pi\in Q_n}\mu(\hat{0},\pi) = \mu(\hat{0},\hat{0}) = 1,$$

and the proof follows.

This proof was suggested by D. Grieser.

5.62. **a.** The case $r=0$ is trivial, so assume $r>0$. Let Γ_A be the corresponding bipartite graph, as defined in Section 5.5. Suppose $(\Gamma_A)_i$ is a connected component of Γ_A with vertex bipartition (X_i, Y_i), where $\#X_i = \#Y_i = j$. Suppose $j \geq 2$. The edges of $(\Gamma_A)_i$ may be chosen as follows. Place a bipartite cycle on the vertices (X_i, Y_i) in $\frac{1}{2}(j-1)!\,j!$ ways (as in the proof of Proposition 5.5.10). Replace some edge e of this cycle with m edges, where $1 \leq m \leq r-1$. Replace each edge at even distance from e also with m edges, while each edge at odd distance is replaced with $r-m$ edges. Thus given (X_i, Y_i), there are $\frac{1}{2}(r-1)(j-1)!\,j!$ choices for $(\Gamma_A)_i$ when $j \geq 2$. When $j=1$ there is only one choice. Hence by the exponential formula for 2-partitions,

$$\begin{aligned}\sum_{n\geq 0} f_r(n)\frac{x^n}{n!^2} &= \exp\left(x+\frac{1}{2}(r-1)\sum_{j\geq 2}(j-1)!\,j!\frac{x^j}{j!^2}\right)\\ &= \exp\left(x+\frac{1}{2}(r-1)[-x+\log(1-x)^{-1}]\right)\\ &= (1-x)^{-\frac{1}{2}(r-1)}\,e^{\frac{1}{2}(3-r)x}.\end{aligned}$$

* **b.** When $r=3$ we obtain

$$\sum_{n\geq 0} f_3(n)\frac{x^n}{n!^2} = \frac{1}{1-x} = \sum_{n\geq 0} x^n,$$

so $f_3(n) = n!^2$. Is there a direct combinatorial proof?

5.63. Let $A = P_1 + P_2 + \cdots + P_{2k}$, and let Γ_A be the bipartite graph corresponding to A, as defined in Section 5.5. Write $\Gamma_m = \Gamma_{P_m}$, so Γ_A is the edge union of $\Gamma_1, \Gamma_2, \ldots, \Gamma_{2k}$. Suppose $(\Gamma_A)_i$ is a connected component of Γ_A with vertex bipartition (X_i, Y_i), where $\#X_i = \#Y_i = j \geq 1$. If $j \geq 2$ then $(\Gamma_A)_i$ is obtained by placing a bipartite cycle on (X_i, Y_i) and then replacing each edge with k edges. This can be done in $\frac{1}{2}(j-1)!\,j!$ ways. Write $E(\Gamma)$ for the multiset of edges of the graph Γ. Then $E(\Gamma_m) \cap E((\Gamma_A)_i)$ consists of j vertex-disjoint edges of $(\Gamma_A)_i$. There are precisely two distinct ways to choose j vertex-disjoint edges of $(\Gamma_A)_i$, and each must occur k times among the sets $E(\Gamma_m) \cap E((\Gamma_A)_i)$, for fixed i and for $1 \leq m \leq 2k$. Hence there are $\binom{2k}{k}$ ways to choose the sets $E(\Gamma_m) \cap E((\Gamma_A)_i)$, $1 \leq m \leq 2k$. Thus for $j \geq 2$ there are

$$\frac{1}{2}(j-1)!\,j!\binom{2k}{k} = (j-1)!\,j!\binom{2k-1}{k}$$

choices for each bipartition (X_i, Y_i) with $\#X_i = \#Y_i = j$. When $j=1$ it is clear that there is only one choice. Hence by the exponential formula for 2-partitions,

$$\begin{aligned}\sum_{n\geq 0} N_k(n)\frac{x^n}{n!^2} &= \exp\left[x+\binom{2k-1}{k}\sum_{j\geq 2}(j-1)!\,j!\frac{x^j}{j!^2}\right]\\ &= \exp\left[x+\binom{2k-1}{k}(-x+\log(1-x)^{-1})\right]\\ &= (1-x)^{-\binom{2k-1}{k}}\exp\left[x\left(1-\binom{2k-1}{k}\right)\right].\end{aligned}$$

5.64. **a.** Let M' be M with its first row multiplied by -1. If k is odd, then $(\det M)^k + (\det M')^k = (\operatorname{per} M)^k + (\operatorname{per} M')^k = 0$, from which it follows that $f_k(n) = g_k(n) = 0$. Now

$$2^{n^2} f_2(n) = \sum_M \left(\sum_{\pi \in \mathfrak{S}_n} \pm m_{1,\pi(1)} \cdots m_{n,\pi(n)} \right)^2$$
$$= \sum_{\pi,\sigma \in \mathfrak{S}_n} (\operatorname{sgn} \pi)(\operatorname{sgn} \sigma) \sum_{i,j} \sum_{m_{ij}=\pm 1} m_{1,\pi(1)} \cdots m_{n,\pi(n)}$$
$$\times m_{1,\sigma(1)} \cdots m_{n,\sigma(n)}.$$

If $\pi \neq \sigma$, say $j = \pi(i) \neq \sigma(i)$, then the inner two sums have a factor $\sum_{m_{ij}=\pm 1} m_{ij} = 0$. Hence

$$2^{n^2} f_2(n) = \sum_{\pi \in \mathfrak{S}_n} (\operatorname{sgn} \pi)^2 \sum_{i,j} \sum_{m_{ij}=\pm 1} (m_{1,\pi(1)} \cdots m_{n,\pi(n)})^2$$
$$= \sum_{\pi \in \mathfrak{S}_n} \sum_{i,j} \sum_{m_{ij}} 1$$
$$= 2^{n^2} n!,$$

so $f_2(n) = n!$. The same argument gives $g_2(n) = n!$, since the factors $(\operatorname{sgn} \pi)$ $(\operatorname{sgn} \sigma)$ above turned out to be irrelevant.

Nyquist, Rice, and Riordan (see reference below) attribute this result (in a somewhat more general form) to R. Fortet, *J. Research Nat. Bur. Standards* **47** (1951), 465–470, though it may have been known earlier. For a connection with Hadamard matrices, see C. R. Johnson and M. Newman, *J. Research Nat. Bur. Standards* **78B** (1974), 167–169, and M. Kac, *Probability and Related Topics in Physical Sciences*, vol. I, Interscience, London/New York, 1959, p. 23.

b. Now we get

$$2^{n^2} f_4(n) = \sum_{\rho,\pi,\sigma,\tau \in \mathfrak{S}_n} (\operatorname{sgn} \rho)(\operatorname{sgn} \pi)(\operatorname{sgn} \sigma)(\operatorname{sgn} \tau)$$
$$\times \sum_{i,j} \sum_{m_{ij}=\pm 1} \prod_{k=1}^{n} m_{k,\rho(k)} m_{k,\pi(k)} m_{k,\sigma(k)} m_{k,\tau(k)}. \tag{5.142}$$

We get a nonzero contribution only when the product P in (5.142) is a perfect square (regarded as a monomial in the variables m_{rs}). Equivalently, if we identify a permutation with its corresponding permutation matrix then $\rho + \pi + \sigma + \tau$ has entries 0, 2, or 4. We claim that in this case the product $\rho\pi\sigma\tau$ is an even permutation. One way to see this is to verify that a fixed cycle C occurs an even number of times (0, 2, or 4, with 4 possible only for singletons) among the four permutations $\rho\rho^{-1}$, $\pi\rho^{-1}$, $\sigma\rho^{-1}$, $\tau\rho^{-1}$. Hence $\rho\rho^{-1}\pi\rho^{-1}\sigma\rho^{-1}\tau\rho^{-1}$ is even, and so also $\rho\pi\sigma\tau$. It follows that the factor $(\operatorname{sgn} \rho)(\operatorname{sgn} \pi)(\operatorname{sgn} \sigma)(\operatorname{sgn} \tau)$ in (5.142) is equal to 1 for all nonzero terms. Hence the right-hand side of (5.142) is equal to $2^{n^2} N_2(n)$, where $N_2(n)$ is the

number of 4-tuples $(\rho, \pi, \sigma, \tau) \in \mathfrak{S}_n^4$ with every entry of $\rho + \pi + \sigma + \tau$ equal to 0, 2, or 4. Hence (5.120) follows from Exercise 5.63.

The computation is identical for $g_4(n)$, since the factor $(\operatorname{sgn}\rho)(\operatorname{sgn}\pi)(\operatorname{sgn}\sigma)(\operatorname{sgn}\tau)$ turned out to be irrelevant.

Equation (5.120) is due to H. Nyquist, S. O. Rice, and J. Riordan, *Quart. J. Appl. Math.* **12** (1954), 97–104, using a different technique. They prove a more general result wherein the matrix entries are identically distributed independent random variables symmetric about 0. The method used here applies equally well to this more general result.

c. It is clear from the above proof technique that $f_{2k}(n) < g_{2k}(n)$ provided there are permutations $\pi_1, \ldots, \pi_{2k} \in \mathfrak{S}_n$ such that $\pi_1 + \cdots + \pi_{2k}$ has even entries and $\pi_1 \cdots \pi_{2k}$ is an odd permutation. For $n = 3$ and $k = 3$ we can take $\{\pi_1, \pi_2, \ldots, \pi_6\} = \mathfrak{S}_3$. For larger values of n and k we can easily construct examples from the example for $n = 3$ and $k = 3$.

d. Let $M \in \mathcal{D}_{n+1}$. Multiply each column of M by ± 1 so that the first row consists of 1's. Multiply each row except the first by ± 1 so that the first column contains -1's in all positions except the first. Now add the first row to all the other rows. The submatrix obtained by deleting the first row and column will be an $n \times n$ matrix $2M'$, where M' is a 0–1 matrix. Expanding by the first column yields $\det M = \pm 2^n(\det M')$. This map $M \longmapsto M'$ produces each $n \times n$ 0–1 matrix the same number (viz., 2^{2n+1}) of times. From this it follows easily that $f_k'(n) = 2^{-kn} f_k(n+1)$ when k is even. When k is odd, one can see easily that $f_k'(n) = 0$.

We leave the easy case of $g_1'(n)$ to the reader and consider $g_2'(n)$. As in (a) or (b), we have

$$2^{n^2} g_2'(n) = \sum_{\pi,\sigma \in \mathfrak{S}_n} \sum_{i,j} \sum_{m_{ij}=0,1} \prod_{k=1}^{n} m_{k,\pi(k)} m_{k,\sigma(k)}.$$

Suppose that the matrix $\pi + \sigma$ has r 2's, and hence $2n - r$ 1's. Equivalently, $\pi\sigma^{-1}$ has r fixed points. Then

$$\sum_{i,j} \sum_{m_{ij}=0,1} \prod_{k=1}^{n} m_{k,\pi(k)} m_{k,\sigma(k)} = 2^{n^2-2n+r}.$$

Since we can choose any $\pi \in \mathfrak{S}_n$ and then choose σ so that $\pi\sigma^{-1}$ has r fixed points, it follows that

$$2^{n^2} g_2'(n) = n! 2^{n^2-2n} \sum_{\pi \in \mathfrak{S}_n} 2^{c_1(\pi)} = n!\, 2^{n^2-2n} h(n),$$

say, where π has $c_1(\pi)$ fixed points. Setting $t_1 = 2$ and $t_2 = t_3 = \cdots = 1$ in (5.30) yields

$$\begin{aligned}\sum_{n \geq 0} h(n) \frac{x^n}{n!} &= \exp\left(2x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots\right) \\ &= \sum_{n \geq 0} n! \left(1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!}\right) \frac{x^n}{n!},\end{aligned}$$

so $h(n) = n!\left(1 + \frac{1}{1!} + \cdots + \frac{1}{n!}\right)$, and $g_2'(n)$ is as claimed. (One could also give a proof using Proposition 5.5.8 instead of (5.30).)

5.65. a. Given a function $g : \mathbb{N} \times \mathbb{N} - \{(0,0)\} \to K$, define a new function $h : \mathbb{N} \times \mathbb{N} \to K$ by

$$h(m, n) = \sum g(\#A_1, \#B_1) \cdots g(\#A_k, \#B_k),$$

where the sum is over all sets $\{(A_1, B_1), \ldots, (A_k, B_k)\}$, where $A_j \subseteq [m]$ and $B_j \subseteq [n]$, satisfying:

(i) For no j do we have $A_j = B_j = \emptyset$.
(ii) The nonempty A_j's form a partition of the set $[m]$.
(iii) The nonempty B_j's form a partition of the set $[n]$.

(Set $h(0, 0) = 1$.) In the same way that Corollary 5.1.6 is proved, we obtain

$$\sum_{m,n \geq 0} h(m, n) \frac{x^m y^n}{m!\, n!} = \exp \sum_{\substack{i,j \geq 0 \\ (i,j) \neq (0,0)}} g(i, j) \frac{x^i y^j}{i!\, j!}.$$

Now let $A = (a_{ij})$ be an $m \times n$ matrix of the type being counted. Let Γ_A be the bipartite graph with vertex bipartition $(\{x_1, \ldots, x_m\}, \{y_1, \ldots, y_n\})$, with a_{ij} edges between x_i and y_j. The connected components $\Gamma_1, \ldots, \Gamma_k$ of Γ_A define a set $\{(A_1, B_1), \ldots, (A_k, B_k)\}$ satisfying (i)–(iii) above, viz., $i \in A_j$ if x_i is a vertex of Γ_j, and $i \in B_j$ if y_i is a vertex of Γ_j. Every connected component of Γ_A must be a path (of length ≥ 0) or a cycle (of even length ≥ 2). We have the following number of possibilities for a component with i vertices among the x_k's and j among the y_k's:

$$\begin{aligned} (i, j) = (0, 1) \text{ or } (1, 0) &: 1 \\ (1, 1) &: 2 \\ (i, i+1) \text{ or } (i+1, i) &: \tfrac{1}{2} i!(i+1)!, \quad i \geq 1 \\ (i, i) &: i!^2 + \tfrac{1}{2}(i-1)!i!, \quad i \geq 2 \\ \text{all others} &: 0. \end{aligned}$$

Hence

$$F(x, y) = \exp\left[x + y + 2xy + \frac{1}{2} \sum_{i \geq 1} (x^i y^{i+1} + x^{i+1} y^i) + \sum_{i \geq 2} \left(1 + \frac{1}{2i}\right) x^i y^i\right],$$

which simplifies to the right-hand side of (5.121).

b. For any power series $G(x, y) = \sum c_{mn} x^m y^n$, let $\mathcal{D}G(x, y) = \sum c_{nn} t^n$. The operator $\mathcal{D}$ preserves infinite linear combinations; and if $G(x, y) =$

$H(x, y, xy)$ for some function H, then $\mathcal{D}G(x, y) = \mathcal{D}H(x, y, t)$. Hence

$$\sum_{n\geq 0} f(n, n)\frac{t^n}{n!^2} = \mathcal{D}F(x, y)$$
$$= (1-t)^{-\frac{1}{2}} e^{t(3-t)/2(1-t)} \mathcal{D}\exp\left[\frac{(x+y)\left(1-\frac{1}{2}t\right)}{1-t}\right].$$

But

$$\mathcal{D}\exp\left[\frac{(x+y)\left(1-\frac{1}{2}t\right)}{1-t}\right] = \mathcal{D}\sum_{n\geq 0}\frac{(x+y)^n}{n!}\left(\frac{1-\frac{1}{2}t}{1-t}\right)^n$$
$$= \sum_{n\geq 0}\binom{2n}{n}\frac{t^n}{(2n)!}\left(\frac{1-\frac{1}{2}t}{1-t}\right)^{2n},$$

and the proof follows.

5.66. **a.** If $r \neq s$ then the matrix $\mathbf{L} - r\mathbf{I}$ has s equal rows and hence has rank at most $r+1$. Thus $\mathbf{L}$ has at least $s-1$ eigenvalues equal to r. If $r = s$ then another r rows of $\mathbf{L} - r\mathbf{I}$ are equal, so $\mathbf{L}$ has at least $r + s - 1$ eigenvalues equal to r.

b. By symmetry, $\mathbf{L}$ has at least $r-1$ eigenvalues equal to s.

c. Since the rows of $\mathbf{L}$ sum to 0, there is at least one 0 eigenvalue. The trace of $\mathbf{L}$ is $2rs$. Since this is the sum of the eigenvalues, the remaining eigenvalue must be $2rs - (s-1)r - (r-1)s = r + s$.

d. By the Matrix-Tree Theorem (Theorem 5.6.8) we have

$$c(K_{rs}) = \frac{1}{r+s}(r+s)r^{s-1}s^{r-1} = r^{s-1}s^{r-1},$$

agreeing with Exercise 5.30.

5.67. For each edge $e = \{i, j\}$ associate an indeterminate $x_{ij} = x_{ji}$. Let $\mathbf{L} = \left(L_{ij}\right)$ be the $n \times n$ matrix

$$L_{ij} = \begin{cases} -x_{ij} & \text{if } i \neq j \\ \sum_{\substack{1\leq k\leq n \\ k\neq i}} x_{ik} & \text{if } i = j, \end{cases}$$

Let $\mathbf{L_0}$ denote $\mathbf{L}$ with the last row and column removed. By the Matrix-Tree Theorem (Theorem 5.6.8), we have

$$\sum_T f(T) = \det \mathbf{L_0}(f),$$

where $\mathbf{L_0}(f)$ is obtained from $\mathbf{L_0}$ by substituting $f(e)$ for x_e. Since the (i, i) entry of $\mathbf{L_0}$ has the form $x_{in}+$ other terms, and since x_{in} appears nowhere else in $\mathbf{L_0}$, it follows that we can replace the (i, i) entry of $\mathbf{L_0}$ with a new indeterminate y_i without affecting the distribution of values of $\det \mathbf{L_0}$. Hence $P_n(q)$ is just the number of invertible $(n-1)\times(n-1)$ symmetric matrices over $\mathbb{F}_q$. For q odd this number was computed by L. Carlitz, *Duke Math. J.* **21** (1954), 123–128 (Thm. 3) as part

of a more general result. A more elementary proof, valid for any q, was later given by J. MacWilliams, *Amer. Math. Monthly* **76** (1969), 152–164 (Thm. 2).

This exercise is related to an unpublished question raised by M. Kontsevich. For further information see R. Stanley, Spanning trees and a conjecture of Kontsevich, preprint, available electronically at http://front.math.ucdavis.edu/math.CO/9806055.

5.68. The argument parallels that of Example 5.6.10. Let V be the vector space of all functions $f : \Gamma \to \mathbb{C}$. Define a linear transformation $\Phi : V \to V$ by

$$(\Phi f)(u) = \sum_{v\in\Gamma} \sigma(v) f(u+v).$$

It is easy to check that the characters $\chi \in \hat{\Gamma}$ are the eigenvectors of Φ, with eigenvalue $\sum_{v\in\Gamma} \sigma(v)\chi(v)$. Moreover, the matrix of Φ with respect to the basis Γ of V is just

$$[\Phi] = \left(\sum_{v\in\Gamma} \sigma(v)\right) \cdot \mathbf{I} - \mathbf{L}(D).$$

Hence the eigenvalues of $\mathbf{L}(D)$ are given by $\sum_{v\in\Gamma} \sigma(v)(1 - \chi(v))$ for $\chi \in \hat{\Gamma}$, and the proof follows from Corollary 5.6.6. Note that Example 5.6.10 corresponds to the case $\Gamma = (\mathbb{Z}/2\mathbb{Z})^n$ and σ given by $\sigma(v) = 1$ if v is a unit vector, while $\sigma(v) = 0$ otherwise.

5.69. **a.** Easily seen that $\tau(D, v) = a_1 a_2 \cdots a_{p-1}$.

b. D is connected and balanced, and the oudegree of vertex v_i is $a_{i-1} + a_i$ (with $a_0 = a_p = 0$). Hence by Theorem 5.6.2,

$$\epsilon(D, e) = a_1 a_2 \cdots a_{p-1} \prod_{i=1}^{p} (a_{i-1} + a_i - 1)!.$$

5.70. The argument is completely parallel to that used to prove Corollary 5.6.15. The digraph D_n becomes the graph with vertex set $[0, d-1]^{n-1}$ and edges $(a_1 a_2 \cdots a_{n-1}, a_2 a_3 \cdots a_n)$, yielding the answer $d!^{d^{n-1}} d^{-n}$. This result seems to have been first obtained in [1].

5.71. First note that the number q of edges of G is given by $W(2) = 2q$ (since G has no loops or multiple edges). Now $2q = dp$ where p is the number of vertices of G, so p is determined as well. It is easy to see that the numbers λ_j satisfying (5.122) for all $\ell \geq 1$ are unique (consider e.g. the generating function $\sum_{\ell\geq 1} W(\ell)x^{\ell}$), and hence by the proof of Corollary 4.7.3 are the nonzero eigenvalues of the adjacency matrix $\mathbf{A}$ of G. Since $\mathbf{A}$ has p eigenvalues in all, it follows that $p - m$ of them are equal to 0. A number of arguments are available to show that the largest eigenvalue λ_1 is equal to d. Since G is regular, the eigenvalues of the Laplacian matrix $\mathbf{L}$ of G are the numbers $d - \lambda_j$, together with the eigenvalue d of multiplicity $p - m$. Hence by the Matrix-Tree Theorem (Theorem 5.6.8),

$$c(G) = \frac{d^{p-m}}{p} \prod_{j=2}^{m} (d - \lambda_j).$$

5.72. There is a standard bijection $T \mapsto T^*$ between the spanning trees T of G and those of G^*, namely, if T has edge set $\{e_1, \ldots, e_r\}$, then T^* has edge set $E^* - \{e_1^*, \ldots, e_r^*\}$, where E^* denotes the edge set of G^*. Hence $c(G) = c(G^*)$. Let $\mathbf{L}_0(G^*)$ denote $\mathbf{L}(G^*)$ with the row and column indexed by the outside vertex deleted. It is easy to see that $\mathbf{L}_0(G^*) = 4\mathbf{I} - \mathbf{A}(G')$, and the proof follows from Theorem 5.6.8.

This result is due to D. Cvetković and I. Gutman, *Publ. Inst. Math. (Beograd)* **29** (1981), 49–52. They give an obvious generalization to planar graphs all of whose bounded regions have the same number of boundary edges. See also D. Cvetković, M. Doob, I. Gutman, and A. Torĝasev, *Recent Results in the Theory of Graph Spectra*, Annals of Discrete Mathematics **36**, North-Holland, Amsterdam, 1988 (Thm. 3.34). For some related work, see T. Chow, *Proc. Amer. Math. Soc.* **125** (1997), 3155–3161; M. Ciucu, *J. Combinatorial Theory (A)* **81** (1998), 34–68; D. E. Knuth, *J. Alg. Combinatorics* **6** (1997), 253–257; and R. Stanley, *Discrete Math.* **157** (1996), 375–388 (Problem 251).

5.74. **a.** Let $\mathbf{J}$ be the $p \times p$ matrix of all 1's. As in the proof of Lemma 5.6.14, we have that $\mathbf{A}^\ell = \mathbf{J}$ and that the eigenvalues of $\mathbf{A}$ are $p^{1/\ell}$ (once) and 0 ($p-1$ times). (Note that since $\operatorname{tr} \mathbf{A}$ is an integer, it follows that $p = r^\ell$ for some $r \in \mathbb{P}$. Part (d) of this exercise gives a more precise result.)

b. The number of loops is $\operatorname{tr} \mathbf{A} = r$, where $p = r^\ell$ as above.

c. Since by hypothesis there is a walk between any two vertices of D, it follows that D is connected. Since $\mathbf{A}$ has a unique eigenvalue equal to r, there is a unique corresponding eigenvector E (up to multiplication by a nonzero scalar). Since E is also an eigenvector of $\mathbf{A}^\ell = \mathbf{J}$ with eigenvalue $r^\ell = p$, it follows that E is the (column) vector of all 1's. The equation $\mathbf{A}E = rE$ shows that every vertex of D has outdegree r. If we take the transpose of both sides of the equation $\mathbf{A}^\ell = \mathbf{J}$, then we get $\left(\mathbf{A}^t\right)^\ell = \mathbf{J}$. Thus the same reasoning shows that $\mathbf{A}^t E = rE$, so every vertex of D has indegree r.

d. The above argument shows that $r = d$ (or $p = d^\ell$).

e. Since every vertex of D has outdegree r, we have $\mathbf{L} = r\mathbf{I} - \mathbf{A}$. Hence by (a) the eigenvalues of $\mathbf{L}$ are r ($p-1$ times) and 0 (once). It follows from Corollary 5.6.7 that

$$\begin{aligned} \epsilon(D, e) &= \frac{1}{p} r^{p-1} (r-1)!^p \\ &= r^{-(\ell+1)} r!^{r^\ell}. \end{aligned}$$

The *total* number of Eulerian tours is just

$$\epsilon(D) = rp \cdot \epsilon(D, e) = r!^{r^\ell}.$$

f. We want to find all $p \times p$ matrices $\mathbf{A}$ of nonnegative integers such that $\mathbf{A}^\ell = \mathbf{J}$. If we ignore the hypothesis that the entries of $\mathbf{A}$ are nonnegative integers, then a simple linear-algebra argument shows that $\mathbf{A} = r^{-\ell+1}\mathbf{J} + \mathbf{N}$ where $\mathbf{N}^\ell = \mathbf{0}$ and $\mathbf{NJ} = \mathbf{JN} = \mathbf{0}$. Equivalently, if $\boldsymbol{e}_i$ denotes the i-th unit coordinate vector, then $\mathbf{N}^\ell = \mathbf{0}$, $\mathbf{N}(\boldsymbol{e}_1 + \cdots + \boldsymbol{e}_p) = 0$, and the space of all vectors $a_1\boldsymbol{e}_1 + \cdots + a_p\boldsymbol{e}_p$ with $\sum a_i = 0$ is $\mathbf{N}$-invariant. Conversely, for

any such $\mathbf{N}$ the matrix $\mathbf{A} = r^{-\ell+1}\mathbf{J} + \mathbf{N}$ satisfies $\mathbf{A}^\ell = \mathbf{J}$. If we choose $\mathbf{N}$ to have integer entries and let c be a large enough integer so that the matrix $\mathbf{B} = c\mathbf{J} + \mathbf{N}$ has nonnegative entries, then $\mathbf{B}$ will be the adjacency matrix of a digraph with the *same* number of paths (not necessarily just one path) of length ℓ between any two vertices. For instance, let $p = 3$ and (writing column vectors as row vectors for simplicity) define $\mathbf{N}$ by $\mathbf{N}[1, 1, 1] = [0, 0, 0]$, $\mathbf{N}[1, -1, 0] = [2, -1, -1]$, and $\mathbf{N}[2, -1, -1] = [0, 0, 0]$. Then

$$2\mathbf{J} + \mathbf{N} = \begin{bmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 4 & 1 & 1 \end{bmatrix},$$

and $(2\mathbf{J} + \mathbf{N})^2 = 12\mathbf{J}$. Hence $2\mathbf{J} + \mathbf{N}$ is the adjacency matrix of a digraph with 12 paths of length two between any two vertices. It is more difficult to obtain a digraph, other than the de Bruijn graphs, with a *unique* path of length ℓ between two vertices, but such examples were given by M. Capalbo and H. Frederickson (independently). The adjacency matrix of Capalbo's example (with a unique path of length two between any two vertices) is the following:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

6

Algebraic, D-Finite, and Noncommutative Generating Functions

6.1 Algebraic Generating Functions

In this chapter we will investigate two classes of generating functions called algebraic and D-finite generating functions. We will also briefly discuss the theory of noncommutative generating functions, especially their connection with rational and algebraic generating functions. The algebraic functions are a natural generalization of rational functions, while D-finite functions are a natural generalization of algebraic functions. Thus we have the hierarchy

$$\begin{array}{c} D\text{-finite} \\ | \\ \text{algebraic} \\ | \\ \text{rational} \end{array} \tag{6.1}$$

Various other classes could be added to the hierarchy, but the three classes of (6.1) seem the most useful for enumerative combinatorics.

6.1.1 Definition. Let K be a field. A formal power series $\eta \in K[[x]]$ is said to be *algebraic* if there exist polynomials $P_0(x), \ldots, P_d(x) \in K[x]$, not all 0, such that

$$P_0(x) + P_1(x)\eta + \cdots + P_d(x)\eta^d = 0. \tag{6.2}$$

The smallest positive integer d for which (6.2) holds is called the *degree* of η.

Note that an algebraic series η has degree one if and only if η is rational. The set of all algebraic power series over K is denoted $K_{\text{alg}}[[x]]$.

6.1.2 Example. Let $\eta = \sum_{n\geq 0} \binom{2n}{n} x^n$. By Exercise 1.4(a) we have $(1-4x)\eta^2 - 1 = 0$. Hence η is algebraic of degree one or two. If K has characteristic 2 then

$\eta = 1$, which has degree one. Otherwise it is easy to see that η has degree two. Namely, if $\deg(\eta) = 1$ then $\eta = P(x)/Q(x)$ for some polynomials $P(x), Q(x) \in K[x]$. Thus

$$(1 - 4x)P(x)^2 = Q(x)^2.$$

The degree (as a polynomial) of the left-hand side is odd while that of the right-hand side is even, a contradiction.

It is a great convenience in dealing with algebraic series to work over fields rather than over the rings $K[x]$ and $K[[x]]$. Thus we review some basic facts about the quotient fields of $K[x]$ and $K[[x]]$. The quotient field of $K[x]$ is just the field $K(x)$ of rational functions over K (since every rational function is by definition a quotient of polynomials). The ring $K[[x]]$ is almost a field; we simply need to invert x to obtain a field. (In other words, $K[[x]]$ is a local domain, and x generates the maximal ideal.) Thus the quotient field of $K[[x]]$ is given by

$$K((x)) = K[[x]][1/x].$$

Every element of $K((x))$ may be regarded as a Laurent series $\eta = \sum_{n \geq n_0} a_n x^n$ for some $n_0 \in \mathbb{Z}$ (depending on η). To see that such Laurent series indeed form a field, the only field axiom that offers any difficulty is the existence of multiplicative inverses. But if $\eta = \sum_{n \geq n_0} a_n x^n$ with $a_{n_0} \neq 0$, then $\eta = x^{n_0}\rho$ where ρ is an ordinary power series with nonzero constant term. Hence $\rho^{-1} \in K[[x]]$, so $\eta^{-1} = x^{-n_0}\rho^{-1} \in K((x))$, as desired.

Recall from any introductory algebra text that if D is an (integral) domain containing the field $K(x)$, then $\eta \in D$ is said to be *algebraic* over $K(x)$ if there exist elements $F_0(x), \ldots, F_d(x) \in K(x)$, not all 0, such that

$$F_0(x) + F_1(x)\eta + \cdots + F_d(x)\eta^d = 0. \tag{6.3}$$

The least such d is the *degree* of η over the field $K(x)$, denoted $\deg_{K(x)}(\eta)$. It is also the dimension of the field $K(x, \eta)$ (obtained by adjoining η to $K(x)$) as a vector space over $K(x)$. Equivalently, η is algebraic over $K(x)$ if and only if the $K(x)$-vector space spanned by $\{1, \eta, \eta^2, \ldots\}$ is finite-dimensional (in which case its dimension is $\deg_{K(x)}(\eta)$). Moreover (again from any introductory algebra text) the set of $\eta \in D$ that are algebraic over $K(x)$ form a subring of D containing $K(x)$ (and hence a $K(x)$-*subalgebra* of D, i.e., a subring of D that is also a vector space over $K(x)$).

Suppose (6.3) holds, and let $P(y) = F_0(x) + F_1(x)y + \cdots + F_d(x)y^d \in K(x)[y]$. Thus $P(y)$ is a polynomial in the indeterminate y with coefficients in the field $K(x)$. We then have $d = \deg_{K(x)}(\eta)$ if and only if the polynomial $P(y)$ is *irreducible*. If we normalize (6.3) by dividing by $F_d(x)$ so that $P(y)$ is monic,

and if $d = \deg_{K(x)}(\eta)$, then the equation (6.3) is *unique*; for otherwise we could subtract two such equations and obtain one of smaller degree.

Note that we can multiply (6.3) by a common denominator of the F_i's and hence can assume that the F_i's are polynomials. Thus we see that $\eta \in K[[x]]$ is algebraic (as defined by Definition 6.1.1) if and only if it is algebraic over $K(x)$. The same is true for $\eta \in K((x))$. The set of all algebraic Laurent series over $K(x)$ is denoted $K_{\text{alg}}((x))$. Hence

$$K_{\text{alg}}[[x]] = K_{\text{alg}}((x)) \cap K[[x]]. \tag{6.4}$$

Again by standard algebraic arguments, $K_{\text{alg}}[[x]]$ is a subring of $K[[x]]$ containing $K(x) \cap K[[x]]$, the ring of rational functions P/Q with $P, Q \in K[x]$ and $Q(0) \neq 0$.

6.1.3 Example. Let us consider Example 6.1.2 from a somewhat more algebraic viewpoint. The series $\eta = \sum_{n\geq 0} \binom{2n}{n} x^n$ will have degree two if and only if $(1-4x)y^2 - 1$ is irreducible as a polynomial in y over $K(x)$. A quadratic polynomial $p = ay^2 + by + c$ is irreducible over a field F of characteristic not equal to two if and only if its discriminant $\text{disc}(p) = b^2 - 4ac$ is not a square in F. Now $\text{disc}((1-4x)y^2 - 1) = 4(1-4x)$, which is not a square, since its degree (as a polynomial in x) is odd. Hence $\deg_{K(x)}(\eta) = 2$ if $\text{char}K \neq 2$.

For most enumerative purposes involving algebraic series it suffices to work with Laurent series $y \in K((x))$. Note, however, that there exist elements η in some extension field of $K(x)$ that are algebraic over $K(x)$ but that cannot be represented as elements of $K((x))$. The simplest such η are defined by $\eta^N = x$ for $N \geq 2$. This suggests that we look at formal series $\eta = \sum_{n\geq n_0} a_n x^{n/N}$, where N is a positive integer depending on η. Such a series is called a *fractional (Laurent) series* (or *Puiseux series*). If we can take $n_0 = 0$ then we have a *fractional power series*. Let $K^{\text{fra}}((x))$ (respectively, $K^{\text{fra}}[[x]]$) denote the ring of all fractional Laurent series (respectively, fractional power series) over K. Thus $K^{\text{fra}}((x)) = K((x))[x^{1/2}, x^{1/3}, x^{1/4}, \ldots]$, i.e., every $\eta \in K^{\text{fra}}((x))$ can be written as a *polynomial* in $x^{1/2}, x^{1/3}, \ldots$ (and hence involving only finitely many of them) with coefficients in $K((x))$; and conversely every such polynomial is a fractional series. Thus for instance $\sum_{N\geq 1} x^{1/N}$ is not a fractional series in our sense of the term. It's easy to see that $K^{\text{fra}}[[x]]$ is a ring, and that $K^{\text{fra}}((x))$ is the quotient field of $K^{\text{fra}}[[x]]$. For instance, the product of $\sum_{m\geq m_0} a_m x^{m/M}$ and $\sum_{n\geq n_0} b_n x^{n/N}$ will be a series of the form $\sum_{n\geq m_0 N + n_0 M} c_n x^{n/MN}$.

For the remainder of this section we will develop some basic properties of fractional series and algebraic series. An understanding of these properties provides some interesting insight into the formal aspects of algebraic series, but such an understanding is not really necessary for solving enumerative problems. The reader may skip the remainder of this section with little loss of continuity. (The only real exception is the proof of Theorem 6.3.3.)

6.1.4 Proposition. *The field $K^{\text{fra}}((x))$ is an algebraic extension of $K((x))$, i.e., every $\eta \in K^{\text{fra}}((x))$ satisfies an equation*

$$P_0(x) + P_1(x)\eta + \cdots + P_d(x)\eta^d = 0,$$

where each $P_i(x) \in K((x))$ and not all $P_i(x) = 0$.

Proof. Let $\eta = \sum_{n \geq n_0} a_n x^{n/N} \in K^{\text{fra}}((x))$. There are then (unique) series $\eta_0, \eta_1, \ldots, \eta_{N-1} \in K((x))$ such that $\eta = \eta_0 + x^{1/N}\eta_1 + x^{2/N}\eta_2 + \cdots + x^{(N-1)/N}\eta_{N-1}$. The series (consisting of a single term each) $x^{1/N}, x^{2/N}, \ldots, x^{(N-1)/N}$ are clearly algebraic over $K((x))$. Since for any extension field E of any field F, the elements of E that are algebraic over F form a subfield of E containing F, we have that η is algebraic over $K((x))$, as desired. □

A considerably deeper result is the following.

6.1.5 Theorem. *Let K be an algebraically closed field of characteristic zero (e.g., $K = \mathbb{C}$). Then the field $K^{\text{fra}}((x))$ is algebraically closed (and hence by Proposition 6.1.4 is an algebraic closure of $K((x))$).*

Theorem 6.1.5 is known as *Puiseux's theorem.* We omit the rather lengthy proof, since we only use Theorem 6.1.5 here to prove the rather peripheral Theorem 6.3.3. For some references to proofs, see the Notes section of this chapter. Somewhat surprisingly, Theorem 6.1.5 is false for algebraically closed fields K of characteristic $p > 0$; see Exercise 6.4.

⋆ For the remainder of this section, unless explicitly stated otherwise, we assume that K is algebraically closed of characteristic zero. ⋆

Since $K(x)$ is a subfield of $K((x))$, it follows from Theorem 6.1.5 that $K^{\text{fra}}((x))$ contains the algebraic closure $\overline{K(x)}$ of $K(x)$. Thus any algebraic power series, as defined in Definition 6.1.1, can be represented as a fractional power series (and any algebraic Laurent series can be represented as a fractional Laurent series).

Suppose that $P(y) \in K((x))[y]$ is an *irreducible* polynomial in y of degree d over the field $K((x))$. The d fractional series $\eta_1, \ldots, \eta_d \in K^{\text{fra}}((x))$ satisfying $P(\eta_j) = 0$ are called *conjugates* of one another. The next result describes the relation between conjugate series.

6.1.6 Proposition. *With $P(y)$ as above, let $\eta = \sum_{n \geq n_0} a_n x^{n/N}$ satisfy $P(\eta) = 0$. Then the least possible value of N is equal to d, and the d conjugates to η are given by*

$$\eta_j = \sum_{n \geq n_0} a_n \zeta^{jn} x^{n/d}, \qquad 0 \leq j < d,$$

where ζ is a primitive d-th root of unity. (Note that ζ exists in K, since K is algebraically closed of characteristic zero.)

NOTE. The intuition behind Proposition 6.1.6 is the following. Let $t_j = \zeta^j x^{1/d}$. Then $t_j^d = x$. Hence each t_j is "just as good" a d-th root of x as $x^{1/d}$ itself and may be substituted for $x^{1/d}$ in η without affecting the validity of the equation $P(\eta) = 0$. A little field theory can make this argument rigorous, though we give a naive proof.

Sketch of the Proof. Let $\eta = \eta(x) = \sum_{n\geq n_0} a_n x^{n/N}$ with N minimal (given η) and let ξ be a primitive N-th root of unity. Let

$$P(y) = F_0(x) + F_1(x)y + \cdots + F_d(x)y^d,$$

so

$$P(\eta) = F_0(x) + F_1(x)\eta + \cdots + F_d(x)\eta^d = 0. \tag{6.5}$$

Now substitute $\xi^j x^{1/N}$ for $x^{1/N}$ in (6.5). Set $\eta_j^* = \sum_{n\geq n_0} a_n \xi^{jn} x^{n/N}$. Each $F_i(x)$ remains unchanged, so $P(\eta_j^*) = 0$, i.e., each η_j^* $(0 \leq j < N)$ is a conjugate of η. Hence $N \leq d$.

On the other hand, let $e_k = e_k(\eta_0^*, \ldots, \eta_{N-1}^*)$ be the k-th elementary symmetric function in $\eta_0^*, \ldots, \eta_{N-1}^*$, for $1 \leq k \leq N$. Each e_k, regarded as a Laurent series in $x^{1/N}$, is invariant under substituting $\xi x^{1/N}$ for $x^{1/N}$. Hence $e_k \in K((x))$, so

$$\prod_{j=0}^{N-1}(y - \eta_j^*) = \sum_{k=0}^{N}(-1)^k e_k y^{N-k} \in K((x))[y].$$

Since $P(y)$ is irreducible, there follows $N \geq d$. Hence $N = d$, and the proof follows. □

The following corollary to Proposition 6.1.6 is sometimes regarded as part of Puiseux's theorem (Theorem 6.1.5).

6.1.7 Corollary. *Suppose that $\eta \in K^{\mathrm{fra}}((x))$ is algebraic of degree d over $K(x)$. Then there exist positive integers $c_1, \ldots, c_r$ satisfying $c_1 + \cdots + c_r = d$, and fractional Laurent series*

$$\eta_j = \sum_{n\geq n_0} a_{j,n} x^{n/c_j}, \quad 1 \leq j \leq r,$$

such that the d conjugates of η (over $K(x)$) are given by

$$\sum_{n\geq n_0} a_{j,n} \zeta_{c_j}^{kn} x^{n/c_j},$$

where $1 \leq j \leq r$, $0 \leq k < c_j$, and ζ_{c_j} denotes a primitive c_j-th root of unity.

Proof. Immediate from Proposition 6.1.6, since conjugates over $K((x))$ remain conjugate over the subfield $K(x)$. □

Now suppose that we are given a polynomial equation

$$P(y) = F_0(x) + F_1(x)y + \cdots + F_d(x)y^d = 0, \tag{6.6}$$

where $F_j(x) \in K((x))$. What information can we read off directly from (6.6) about the d roots $\eta_1, \ldots, \eta_d$ of $P(y)$? For instance, how many roots are ordinary power series (elements of $K[[x]]$)? In general there is no simple way to obtain such information, but there are some useful sufficient conditions for various properties of the η_j's.

6.1.8 Proposition. *Let*

$$P(y) = F_0(x) + F_1(x)y + \cdots + F_d(x)y^d,$$

where each $F_i(x) \in K[[x]]$ (or even $K^{\mathrm{fra}}[[x]]$) and for at least one j we have $F_j(0) \neq 0$. Let m be the number of solutions $y = \eta_i$ to $P(y) = 0$ that are fractional power series (i.e., no negative exponents). Then

$$m = \max\{j : F_j(0) \neq 0\}.$$

Proof. For any fractional series η, let

$$\nu(\eta) = \min\{a : [x^a]\eta \neq 0\},$$

and call $\nu(\eta)$ the *x-degree* of η (to distinguish it from the degree of η as an algebraic element over $K((x))$). Let $P(y) = F_d(x)(y-\eta_1)\cdots(y-\eta_d)$, where $\nu(\eta_i) = a_i < 0$ for $1 \leq i \leq d-m$, and $\nu(\eta_i) \geq 0$ for $d-m+1 \leq i \leq d$. Consider the elementary symmetric function

$$e_{d-m} = e_{d-m}(\eta_1, \ldots, \eta_d) = \sum_{i_1 < \cdots < i_{d-m}} \eta_{i_1} \cdots \eta_{i_{d-m}}.$$

Thus $[y^m]P(y) = (-1)^{d-m} F_d(x) e_{d-m}$. Now e_{d-m} has a term $\eta_1\eta_2\cdots\eta_{d-m}$ of x-degree $\sum_{i=1}^{d-m} a_i := A$, while all other terms have greater x-degree. Hence $\nu(e_{d-m}) = A$ exactly. Thus in order for $[y^m]P(y) \in K^{\mathrm{fra}}[[x]]$, we must have $\nu(F_d(x)) \geq -A$. Now every term $\eta_{i_1}\cdots\eta_{i_k}$ of every elementary symmetric function $e_k(\eta_1, \ldots, \eta_d)$ has x-degree at least A. Since some $F_j(0) \neq 0$, it follows that $\nu(F_d(x)) \leq -A$, so $\nu(F_d(x)) = -A$. Hence $F_m(0) \neq 0$. Moreover, if $k < d-m$ then every term of $e_k(\eta_1, \ldots, \eta_d)$ has x-degree strictly greater than A, so $F_d(x)e_k$ (which equals $(-1)^k[y^{d-k}]P(y)$) has strictly positive x-degree, i.e., $F_{d-k}(0) = 0$. Hence $m = \max\{j : F_j(0) \neq 0\}$, as desired. □

If $P(y) = c_0 + c_1 y + \cdots + c_d y^d$ is a polynomial of degree d over a field F with roots $\alpha_1, \ldots, \alpha_d$ in some extension field, then define the *discriminant* of P by

$$\operatorname{disc}(P) = c_d^{2d-2} \prod_{i<j} (\alpha_i - \alpha_j)^2.$$

It is well known that $\operatorname{disc}(P)$ can be expressed as a polynomial in the coefficients c_i of $P(y)$. If we set $\deg c_i = 1$, then $\operatorname{disc}(P)$ is homogeneous of degree $2d-2$; while if we set $\deg c_i = d-i$, then $\operatorname{disc}(P)$ is homogeneous of degree $d(d-1)$. For instance,

$$\begin{aligned}
\operatorname{disc}(ay^2+by+c) &= b^2-4ac \\
\operatorname{disc}(ay^3+by^2+cy+d) &= -27a^2d^2+18abcd-4ac^3-4b^3d+b^2c^2 \\
\operatorname{disc}(ay^4+by^3+cy^2+dy+e) &= 256a^3e^3-128a^2c^2e^2+144a^2cd^2e \\
&\quad -192a^2bde^2-27a^2d^4+16ac^2e-4ac^3d^2 \\
&\quad -80abc^2de+144ab^2ce^2+18abcd^3 \\
&\quad -6ab^2d^2e-4b^2c^3e+b^2c^2d^2+18b^3cde \\
&\quad -27b^4e^2-4b^3d^3 \\
\operatorname{disc}(ay^d+by+c) &= (-1)^{\binom{d}{2}}a^{d-2}[d^d ac^{d-1}+(-1)^{d-1}(d-1)^{d-1}b^d].
\end{aligned} \tag{6.7}$$

For a proof of (6.7) see Exercise 6.8(a).

6.1.9 Proposition. *Let*

$$P(y) = F_0(x) + F_1(x)y + \cdots + F_d(x)y^d,$$

where each $F_j(x) \in K[[x]]$ and $F_d(0) \neq 0$ (so by Proposition 6.1.8, all the roots of $P(y)$ are fractional power series). Suppose that

$$\operatorname{disc}(P(y))|_{x=0} \neq 0. \tag{6.8}$$

Then every root $\eta_i \in K^{\mathrm{fra}}[[x]]$ of $P(y)$ is an ordinary power series, i.e., $\eta_i \in K[[x]]$.

Proof. Suppose that $P(y)$ has a root $\eta_1 = \sum_{n\geq 0} a_n x^{n/N}$, for some $N > 1$, that is not a power series. If ζ is a primitive N-th root of unity, then by Proposition 6.1.6 (after factoring $P(y)$ into irreducibles) we see that another root of $P(y)$ is $\eta_2 = \sum_{n\geq 0} a_n \zeta^n x^{n/N}$ and that $\eta_1 \neq \eta_2$. Since $\eta_1(0) = \eta_2(0) = a_0$, the difference $\eta_1 - \eta_2$ vanishes at $x = 0$, and hence so does $\operatorname{disc}(P(y))$, contradicting (6.8). □

6.1.10 Example. (a) Let $P(y) = y^2-(x+1)$. Then $F_2(0) = 1 \neq 0$, $\operatorname{disc}(P(y)) = 4(x+1)$, and $\operatorname{disc}(P(y))|_{x=0} = 4 \neq 0$. Thus by Propositions 6.1.8 and 6.1.9, the

equation $P(y) = 0$ has two power series solutions. Indeed, the solutions are given by

$$\pm(1+x)^{1/2} = \pm \sum_{n\geq 0} \binom{\frac{1}{2}}{n} x^n.$$

(b) Let $P(y) = y^2 - x$. Then $F_2(0) \neq 0$, so the two roots are fractional power series by Proposition 6.1.8. Since $\mathrm{disc}(P(y))|_{x=0} = 0$, we cannot tell just from $\mathrm{disc}(P(y))$ whether or not the two roots are in fact ordinary power series. Of course by inspection we see that the roots are $\pm x^{1/2}$, which are not ordinary power series.
(c) Let $P(y) = y^2 - x^2(x+1)$. Again $F_2(0) \neq 0$ and $\mathrm{disc}(P(y))|_{x=0} = 0$. This time, however, the roots *are* ordinary power series.
(d) Let $P(y) = xy^2 - y - 1$. By Proposition 6.1.8 exactly one of the roots is a fractional power series. Proposition 6.1.9 cannot be directly applied, since $F_2(0) = 0$. There are several ways, however, to see that both roots are ordinary Laurent series.

- Suppose that some root was not a Laurent series. Since $d = 2$, the roots η_1 and η_2 are of the form $\sum a_n x^{n/2}$ and $\sum (-1)^n a_n x^{n/2}$. Hence either both or none of the roots are fractional power series. Since we have seen that exactly one root is a fractional power series, it follows that both roots are ordinary Laurent series. More generally, if $r = 1$ (so $c_1 = d$) in Corollary 6.1.7, then either all or none of the η_j are fractional power series.
- Let $z = yx$. Then $xP(y) = z^2 - z - x$. Now Proposition 6.1.9 does apply, and we see that both roots ρ_1, ρ_2 of $z^2 - z - x$ are ordinary power series. Hence the roots $\eta_i = x^{-1}\rho_i$ of $P(y)$ are ordinary Laurent series.
- Since $P(y)$ is quadratic, we can just use the quadratic formula to obtain the two roots

$$\frac{1 \pm \sqrt{1+4x}}{2x} \tag{6.9}$$

where $\sqrt{1+4x} = \sum_{n\geq 0} \binom{1/2}{n} 4^n x^n = 1 + 2x + \cdots$. Hence the plus sign in (6.9) produces a Laurent series η_1 with $\nu(\eta_1) = -1$, while the minus sign yields an ordinary power series η_2.

We mentioned earlier the standard result that the set $K_{\mathrm{alg}}[[x]]$ of algebraic power series forms a subalgebra of $K[[x]]$. Thus if $u, v \in K_{\mathrm{alg}}[[x]]$ and $\alpha, \beta \in K$, then $\alpha u + \beta v$, $uv \in K_{\mathrm{alg}}[[x]]$. A further operation of combinatorial interest that can be performed on power series u and v is the *Hadamard product* $u * v$, defined in Section 4.2. Recall that if $u = \sum f(n)x^n$ and $v = \sum g(n)x^n$, then $u * v := \sum f(n)g(n)x^n$. What effect does this operation have on algebraicity? One can show that if u and v are algebraic, then $u * v$ need not be algebraic. An example

(which takes some work to verify – see Exercise 6.3) is $u = v = (1-4x)^{-1/2} = \sum \binom{2n}{n} x^n$. The next result shows, however, that a weaker result is true.

6.1.11 Proposition. *Let K be a field of characteristic 0. If $u \in K[[x]]$ is algebraic and $v \in K[[x]]$ is rational, then $u * v$ is algebraic.*

Proof. Let $u = \sum f(n)x^n$, $v = \sum g(n)x^n$. By Theorem 4.1.1, there exist $\gamma_1, \dots, \gamma_r \in K$ (since we are assuming K is algebraically closed, though we could just as easily let K be any field of characteristic 0 and work over an algebraic closure of K) and $P_1, \dots, P_r \in K[x]$ such that

$$g(n) = \sum P_i(n)\gamma_i^n, \qquad n \gg 0. \tag{6.10}$$

Now changing finitely many coefficients of v has no effect on whether $u * v$ is algebraic [why?], so we can assume that (6.10) holds for *all* $n \geq 0$. Since linear combinations of algebraic functions are algebraic (because $K_{\mathrm{alg}}[[x]]$ is a subalgebra of $K[[x]]$), it suffices to show that for $0 \neq \gamma \in K$, the two series $u_1 = \sum \gamma^n f(n)x^n$ and $u_2 = \sum nf(n)x^n$ are algebraic. Let $P(x, u) = 0$, where $0 \neq P(x, y) \in K[x, y]$. Then $P(\gamma x, y) \neq 0$ and $P(\gamma x, u_1) = 0$, so u_1 is algebraic. Now note that

$$0 = \frac{d}{dx}P(x, u) = \left.\frac{\partial P(x, y)}{\partial x}\right|_{y=u} + u'\left.\frac{\partial P(x, y)}{\partial y}\right|_{y=u}. \tag{6.11}$$

If we assume that we have chosen $P(x, y)$ to be of minimal degree in y (so it is irreducible over $K(x)$), then $\partial P(x, y)/\partial y$ is a nonzero (since $\operatorname{char} K = 0$) polynomial in y of smaller degree than P, so $\left.\frac{\partial P(x,y)}{\partial y}\right|_{y=u} \neq 0$. It therefore follows from (6.11) that

$$u' = -\frac{\left.\frac{\partial P(x,y)}{\partial x}\right|_{y=u}}{\left.\frac{\partial P(x,y)}{\partial y}\right|_{y=u}} \in K(x, u). \tag{6.12}$$

In other words, u' is a rational function of x and u, so u' is algebraic. Hence $u_2 = xu'$ is algebraic, completing the proof. □

We conclude this section with a simple result on algebraic functions that we will need later (see the proof of Theorem 6.7.1). We will be considering Laurent series

$$y = f(x_1, \dots, x_k) \in K(x)((x_1, \dots, x_k)), \tag{6.13}$$

i.e., y is a Laurent series in $x_1, \dots, x_k$ whose coefficients are rational functions (which we regard as Laurent series) in x over K. We will assume that $f(1, \dots, 1)$ is a well-defined element of $K((x))$. More precisely, if $y = \sum_\alpha c_\alpha(x) x_1^{\alpha_1} \cdots x_k^{\alpha_k}$

then $\sum_\alpha c_\alpha(x)$ converges to an element of $K((x))$ in the topology of Section 1.1. For instance, if

$$f(x_1, x_2) = \sum_{i,j \geq 0} x^{i+j} x_1^i x_2^j,$$

then $f(1, 1)$ is the well-defined series $\sum_{n \geq 0}(n + 1)x^n$. On the other hand, if

$$f(x_1, x_2) = \sum_{i,j \geq 0} x^{i-j} x_1^i x_2^j,$$

then $f(1, 1)$ is undefined (e.g., the coefficient of x^0 is the meaningless expression $\sum_{i \geq 0} 1$).

6.1.12 Proposition. *Let $y = f(x_1, \ldots, x_k)$ be as in (6.13). Suppose that $f(1, \ldots, 1)$ is well defined and that y is algebraic over $K(x)(x_1, \ldots, x_k)$. Then $f(1, \ldots, 1)$ is algebraic over $K(x)$.*

Proof. Suppose

$$P_d(x_1, \ldots, x_k)y^d + \cdots + P_0(x_1, \ldots, x_k) = 0, \tag{6.14}$$

where $P_i \in K(x)[x_1, \ldots, x_k]$, the P_i's are relatively prime (as polynomials in $x_1, \ldots, x_k$), and $P_d \neq 0$. By clearing denominators we may assume in fact $P_i \in K[x][x_1, \ldots, x_k] = K[x, x_1, \ldots, x_k]$. We can't simply substitute $x_1 = \cdots = x_k = 1$ in (6.14), since we conceivably might have $P_i(1, \ldots, 1) = 0$ for all i. So instead first set $x_k = 1$. If $P_i(x_1, \ldots, x_{k-1}, 1) = 0$ then $P_i(x_1, \ldots, x_k)$ is divisible by $x_k - 1$. Since the P_i's are relatively prime, some $P_i(x_1, \ldots, x_{k-1}, 1) \neq 0$. Hence $f(x_1, \ldots, x_{k-1}, 1)$ is algebraic over $K(x)(x_1, \ldots, x_{k-1})$. The proof follows by induction on k (the case $k = 0$ being trivial). □

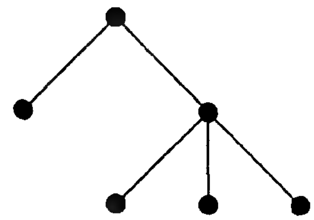
6.2 Examples of Algebraic Series

In this section we will consider some examples of algebraic series that arise in enumerative problems. For some further examples, see Sections 6.3 and 6.7.

We turn first to (unlabeled) plane trees and plane forests, as considered in Section 5.3. Let S be any subset of $\mathbb{P}$, and define a *plane S-tree* to be a plane tree for which any non-endpoint vertex has degree (number of successors) in S. For instance, a plane 2-tree (short for $\{2\}$-tree) is a plane binary tree. There are many interesting combinatorial structures that are "equivalent to" (can be put into a simple one-to-one correspondence with) plane trees. We collect some of the most important of these structures in the next result. We include an example of each structure with $S = \{2, 3\}$, $n = 6$, $m = 4$.

6.2.1 Proposition. *Let $S \subseteq \mathbb{P}$, $n \in \mathbb{P}$, $m \in \mathbb{P}$. There are "nice" bijections between the following sets:*

(i) *plane S-trees with n vertices and m endpoints:*



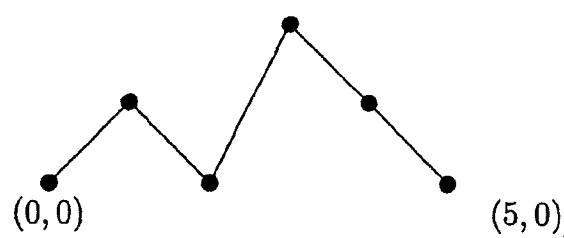
(ii) *sequences $i_1 i_2 \cdots i_{n-1}$, where each $i_j + 1 \in S$ or $i_j = -1$, such that there are a total of $m-1$ values of j for which $i_j = -1$, and such that $i_1 + i_2 + \cdots + i_j \geq 0$ for all j, and $i_1 + i_2 + \cdots + i_{n-1} = 0$:*

$$1, -1, 2, -1, -1;$$

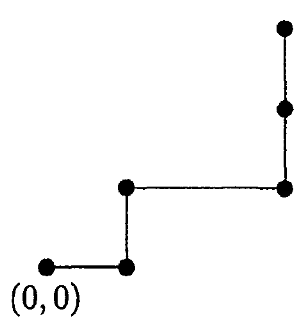
(iii) *parenthesizations (or* bracketings*) of a word of length m subject to $n - m$ k-ary operations, where $k \in S$:*

$$(x(xxx)) \qquad (\textit{one 3-ary and one 2-ary operation});$$

(iv) *paths P in the (x, y) plane from $(0, 0)$ to $(n-1, 0)$ using steps $(1, k)$, where $k + 1 \in S$ or $k = -1$, with a total of $m - 1$ steps of the form $(1, -1)$, such that P never passes below the x-axis:*

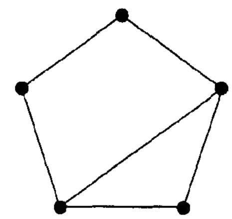


(v) *paths P in the (x, y) plane from $(0, 0)$ to $(m-1, m-1)$, using steps $(k, 0)$ or $(0, 1)$ with $k + 1 \in S$, with a total of $n - 1$ steps, such that P never passes above the line $x = y$:*



(vi) *dissections of a convex $(m+1)$-gon C into $n - m$ regions, each a k-gon with $k - 1 \in S$, by drawing diagonals (necessarily $n - m - 1$ of them) that*

don't intersect in their interiors (when $k = 2$, we must draw two "curved" diagonals between the same pair of vertices of C):



Proof. The bijection between (i) and (ii) was described in Section 5.3 (see Lemma 5.3.9 and the discussion preceding it). Recall that the bijection is obtained by doing a depth-first search through the plane tree τ and recording the integer $(\deg v) - 1$ whenever a vertex v is encountered for the first time. Here we should ignore the last vertex (which will be an endpoint), though in Section 5.3 it was included. For example, the tree of Figure 5-14 gives rise to the sequence

$$2, 0, 1, -1, -1, -1, 1, -1, 1, -1.$$

Now let τ be a plane tree. If τ is just a single vertex, then define the corresponding parenthesization $T_\tau = x$ (a single letter with no operation). Otherwise let the subtrees of the root of τ be $\tau_1, \dots, \tau_j$ (in the given order), and define inductively $W_\tau = (W_{\tau_1} W_{\tau_2} \cdots W_{\tau_j})$. Clearly this construction sets up a bijection between (i) and (iii).

Now given a sequence $i_1 i_2 \cdots i_{n-1}$ enumerated by (ii), let P be the path from $(0, 0)$ to $(n-1, 0)$ with successive steps $(1, i_1), (1, i_2), \dots, (1, i_{n-1})$. This yields a bijection between (ii) and (iv).

The paths of type (v) are simply linear transformations of those of type (iv). More precisely, the sequence $i_1 i_2 \cdots i_{n-1}$ of (ii) corresponds to the path from $(0, 0)$ to $(m-1, m-1)$ whose j-th step is $(k, 0)$ if $i_j = k > 0$ and is $(0, 1)$ if $i_j = -1$.

Finally, consider a convex $(m+1)$-gon C. Fix an edge e_0 of C, called the *root edge*. (The bijection to be described depends on the choice of e_0.) Given a dissection D of C as in (vi), define a plane tree $\tau = \tau(D, e_0)$ as follows. The vertices v_e of τ correspond to the edges e of D. The root vertex of τ is v_{e_0}. When we remove e_0 from D we obtain a sequence (in counterclockwise order) of edge-disjoint dissections $D_1, \dots, D_k$ of polygons $C_1, \dots, C_k$ (where $k+1$ is the number of edges of the region R_0 of D that contains e_0). Let e_i be the edge of D_i (or C_i) that is also an edge of R_0. Define the subtrees $\tau_1, \dots, \tau_k$ of the root v_{e_0} of τ by $\tau_i = \tau(D_i, e_i)$. This gives an inductive definition of τ and establishes a bijection between (vi) and (i). □

To understand this last bijection between (i) and (vi), "one picture is worth a thousand words." Figure 6-1 should make the bijection clear. The edges of D are solid, and the edges of τ are broken. The vertices of D are dots, and of τ are asterisks.

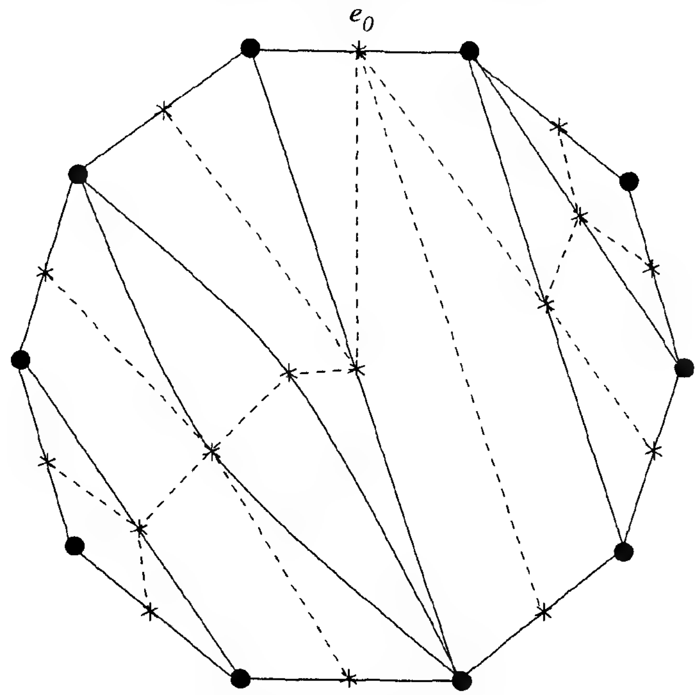


Figure 6-1. A plane tree obtained from a dissected polygon.

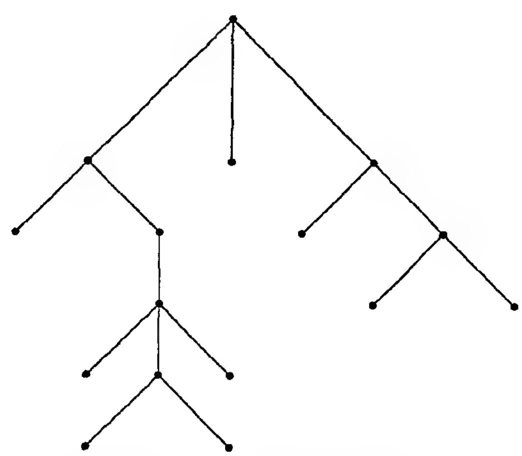


Figure 6-2. The tree of Figure 6-1.

The tree $\tau = \tau(D, e_0)$ is redrawn in Figure 6-2 for greater clarity. Sometimes it is clearer to regard the edges of τ as crossing the nonroot edges of D. Figure 6-3 shows Figure 6-2 redrawn in this way (with e_0 removed for even greater clarity).

One special case of Proposition 6.2.1 of particular interest occurs when S contains a single element $k \geq 2$. In this case the objects discussed in Proposition 6.2.1 exist only when

$$n = kj + 1, \qquad m = (k-1)j + 1$$

for some $j \geq 0$, or equivalently $(k-1)(n-1) = k(m-1)$. Let us write $T_S(m, n)$

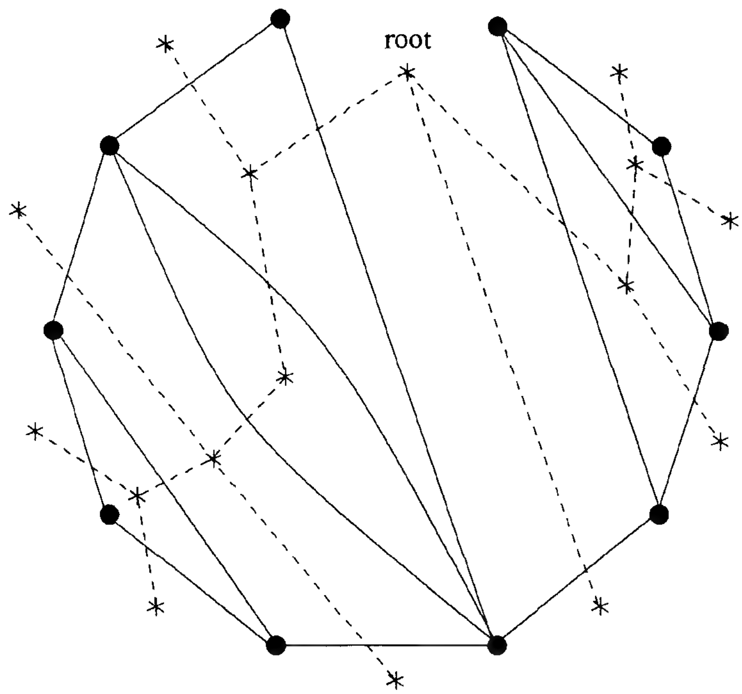


Figure 6-3. Figure 6-1 redrawn.

for the number of plane S-trees with n vertices and m endpoints, and $T_S(n)$ for the total number of plane S-trees with n vertices. We abbreviate $T_{\{k\}}$ by T_k. A special case of Theorem 5.3.10 is the following result.

6.2.2 Proposition. *We have*

$$T_k(m,n) = \begin{cases} \dfrac{1}{n}\dbinom{n}{j} & \text{if } n = kj+1 \text{ and } m = (k-1)j+1 \\ 0 & \text{otherwise.} \end{cases}$$

When $k = 2$, we recover the result (Example 5.3.12) that the number $T_2(n+1, 2n+1) = T_2(2n+1)$ of plane binary trees with $n+1$ endpoints (equivalently, $2n+1$ vertices) is the *Catalan number*

$$C_n := \frac{1}{2n+1}\binom{2n+1}{n} = \frac{1}{n+1}\binom{2n}{n}.$$

The Catalan numbers form one of the most ubiquitous and fascinating sequences of enumerative combinatorics. Proposition 6.2.1 yields a number of combinatorial interpretations of Catalan numbers, all closely connected (since there are simple bijections between the six classes). Let us reiterate the case $S = \{2\}$ of Proposition 6.2.1 directly in terms of Catalan numbers.

6.2.3 Corollary. *The Catalan number C_n counts the following:*

(i) *Plane binary trees with $n+1$ endpoints (or $2n+1$ vertices).*

(ii) *Sequences $i_1 i_2 \cdots i_{2n}$ of 1's and -1's with $i_1 + i_2 + \cdots + i_j \geq 0$ for all j and $i_1 + i_2 + \cdots + i_{2n} = 0$. Such sequences (as well as certain generalizations) are called* ballot sequences, *for the following reason. Suppose that an election is being held between two candidates A and B, and that a 1 (respectively, -1) indicates a vote for A (respectively, B). Then a ballot sequence corresponds to a sequence of $2n$ votes such that A never trails B and the election ends in a tie. (For a generalization of ballot sequences to any number of candidates, see Proposition 7.10.3.)*

(iii) *Ways to parenthesize a string of length $n+1$ subject to a nonassociative binary operation.*

(iv) *Paths P in the (x, y) plane from $(0,0)$ to $(2n, 0)$, with steps $(1,1)$ and $(1,-1)$, that never pass below the x-axis. Such paths are called* Dyck paths.

(v) *Paths P in the (x, y) plane from $(0,0)$ to (n,n), with steps $(1,0)$ and $(0,1)$, that never pass above the line $y = x$.*

(vi) *Ways to dissect a convex $(n+2)$-gon into n triangles by drawing $n-1$ diagonals, no two of which intersect in their interior. Such dissections may be called* triangulations *of an $(n+2)$-gon (with no new vertices).*

A host of other appearances of Catalan numbers in enumerative combinatorics and in other areas of mathematics are given in Exercises 6.19–6.36. Some of the enumerative properties of Catalan numbers are quite surprising and subtle, and do not yield to "transparent" bijections such as those used to establish Corollary 6.2.3.

We are now ready to discuss the connection between plane S-trees and algebraic functions.

6.2.4 Proposition. *Let $S \subseteq \mathbb{P}$, and define*

$$u = u(t, x) = \sum_{n \geq 0} \sum_{m \geq 0} T_S(m, n) t^m x^n.$$

Then

$$u = tx + x \sum_{j \in S} u^j. \tag{6.15}$$

Proof. Note that u^j is the generating function for ordered j-tuples of plane S-trees, i.e.,

$$u^j = \sum_{n \geq 0} \sum_{m \geq 0} T_{S,j}(m, n) t^m x^n,$$

where $T_{S,j}(m, n)$ is the number of ordered j-tuples $(\tau_1, \ldots, \tau_j)$ of plane S-trees

with m endpoints (where by definition an endpoint of $(\tau_1, \ldots, \tau_j)$ is an endpoint of some τ_i) and n vertices.

We obtain a plane S-tree τ from $(\tau_1, \ldots, \tau_j)$ for $j \in S$ by letting $\tau_1, \ldots, \tau_j$ be the subtrees of the root. The number $p(\tau)$ of vertices of τ satisfies $p(\tau) = 1 + \sum p(\tau_i)$, while the number $q(\tau)$ of endpoints satisfies $q(\tau) = \sum q(\tau_i)$. Every plane S-tree τ with more than one point corresponds uniquely to such a j-tuple $(\tau_1, \ldots, \tau_j)$, so (6.15) follows. □

By specializing (6.15) in various ways we can obtain algebraic generating functions. In particular, if v enumerates plane S-trees by number of endpoints (where we must assume $1 \notin S$), then $v = u(t, 1)$ so

$$v = t + \sum_{j \in S} v^j. \tag{6.16}$$

Similarly if w enumerates plane S-trees by number of vertices, then $w = u(1, x)$ so

$$w = x + x \sum_{j \in S} w^j. \tag{6.17}$$

Clearly v and w will be algebraic for suitable choices of S. It can be shown that the following five conditions on S are equivalent when $\operatorname{char} K = 0$:

(i) v is algebraic (assuming $1 \notin S$, so v is defined).
(ii) w is algebraic.
(iii) $\sum_{j \in S} x^j$ is rational.
(iv) S differs by a finite set from a finite union of (infinite) arithmetic progressions in $\mathbb{P}$.
(v) The function $\chi_S : \mathbb{P} \to \{0, 1\}$ defined by $\chi_S(j) = 1$ if $j \in S$ and $\chi_S(j) = 0$ if $j \notin S$ is eventually periodic.

It's easy to see that

$$\text{(i)} \Leftrightarrow \text{(ii)} \Leftarrow \text{(iii)} \Leftarrow \text{(iv)} \Leftrightarrow \text{(v)}.$$

The implications (ii) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (iv) take considerably more work. (They can be deduced from Exercise 4.19(b), (c).)

Note that since (6.16) and (6.17) can be solved explicitly for t and x, respectively, we obtain explicit expressions for the compositional inverses $v^{\langle -1 \rangle}$ and $w^{\langle -1 \rangle}$:

$$v^{\langle -1 \rangle} = t - \sum_{j \in S} t^j, \qquad 1 \notin S$$

$$w^{\langle -1 \rangle} = \frac{x}{1 + \sum_{j \in S} x^j}.$$

6.2.5 Example. Let $S = 4\mathbb{P} \cup (5\mathbb{N}+1) \cup \{9, 11\} - \{1, 12, 16\}$. Then

$$\sum_{j\in S} x^j = \frac{x^4}{1-x^4} + \frac{x}{1-x^5} - \frac{x^{16}}{1-x^{20}} + x^9 + x^{11} - x - x^{12} - x^{16},$$

which is rational. Thus

$$v = x + \frac{v^4}{1-v^4} + \frac{v}{1-v^5} - \frac{v^{16}}{1-v^{20}} + v^9 + v^{11} - v - v^{12} - v^{16},$$

which clearly yields a polynomial equation (of degree 36) satisfied by v.

6.2.6 Example. Let $S = \{k\}$, $k \geq 2$. Combining Proposition 6.2.2 and equation (6.16) yields the following result: Let

$$\begin{aligned} v &= \sum_{n\geq 0} \frac{1}{kn+1}\binom{kn+1}{n} x^{(k-1)n+1} \\ &= \sum_{n\geq 0} \frac{1}{(k-1)n+1}\binom{kn}{n} x^{(k-1)n+1}. \end{aligned} \tag{6.18}$$

Then $v = x + v^k$, so $v = (x - x^k)^{\langle -1 \rangle}$. Of course we can obtain this result directly from the Lagrange inversion formula (Theorem 5.4.1). In fact, Proposition 6.2.2 is a special case of the basis for one of our combinatorial proofs of Lagrange inversion.

Suppose (continuing to assume $\operatorname{char} K = 0$) we let k be *any* element of $K - \{1\}$ (or an indeterminate over K), and define

$$\begin{aligned} y = y(x) &= \sum_{n\geq 0} \frac{1}{kn+1}\binom{kn+1}{n} x^n \\ &= \sum_{n\geq 0} \frac{1}{(k-1)n+1}\binom{kn}{n} x^n. \end{aligned}$$

If $k \in \mathbb{P}$ and v is as in (6.18), then $v = xy(x^{k-1})$. Hence $xy(x^{k-1}) = x + x^k y(x^{k-1})^k$, so

$$y = 1 + xy^k. \tag{6.19}$$

In fact, (6.19) remains valid for any $k \in K$. There are two ways to see this: (a) Lagrange inversion, and (b) equating coefficients of x^n on both sides of (6.19)

yields a tentative polynomial identity involving k. Since we know it's true for all $k \in \mathbb{P} - \{1\}$, it follows that it's valid as a polynomial identity.

6.2.7 Example. An interesting series related to (6.18) is given by

$$z = \sum_{n \geq 0} \binom{kn}{n} x^n. \tag{6.20}$$

Here we may assume $k \in \mathbb{P}$, or even, as in the previous paragraph, $k \in K$. Thus, with v as in (6.18), we have $z(x^{k-1}) = v'$ (where $v' = dv/dx$). Differentiating $v = x + v^k$ yields $v' = 1 + kv'v^{k-1}$. Multiply by v to get $v'v = v + kv'(v - x)$. Solving for v yields

$$v = \frac{kv'x}{1 + (k-1)v'}.$$

Hence from $v = x + v^k$ we get

$$\frac{kv'x}{1 + (k-1)v'} = x + \left(\frac{kv'x}{1 + (k-1)v'} \right)^k,$$

so

$$\frac{v' - 1}{1 + (k-1)v'} = x^{k-1} \left(\frac{kv'}{1 + (k-1)v'} \right)^k.$$

Therefore

$$\frac{z - 1}{1 + (k-1)z} = x \left(\frac{kz}{1 + (k-1)z} \right)^k. \tag{6.21}$$

Hence if $k \in \mathbb{Q}$, then z is algebraic. For further information about the series z in the case $k \in \mathbb{P}$, see Exercise 6.13. Some series that seem closely related to (6.20), such as $\sum \binom{3n}{n,n,n} x^n$ and $\sum \binom{2n}{n}^2 x^n$, can be shown to be nonalgebraic. See Exercise 6.3.

6.2.8 Example. Some interesting enumerative problems correspond to the cases $S = \mathbb{P}$ and $S = \mathbb{P} - \{1\}$. For instance, if $S = \mathbb{P} - \{1\}$ then $v = u(t, 1)$ is the generating function for the total number of dissections of an $(n+2)$-gon C by diagonals not intersecting in the interior of C (and only allowing straight diagonals, so no two-sided regions are formed). Similarly, if $S = \mathbb{P}$ then $w = u(1, x)$ is the generating function for the total number $T_{\mathbb{P}}(n)$ of plane trees with n vertices. From

(6.15), (6.16), and (6.17) we compute:

$$
\begin{aligned}
S = \mathbb{P} - \{1\} \Rightarrow u &= \frac{1 + tx - \sqrt{1 - 2tx - 4tx^2 + t^2x^2}}{2(1+x)} \\
v &= \frac{1 + t - \sqrt{1 - 6t + t^2}}{4} \\
w &= \frac{1 + x - \sqrt{1 - 2x - 3x^2}}{2(1+x)} \\
&= \frac{1}{2}\left(1 - \sqrt{\frac{1-3x}{1+x}}\right)
\end{aligned}
\tag{6.22}
$$

$$
\begin{aligned}
S = \mathbb{P} \Rightarrow u &= \frac{1 - x + tx - \sqrt{1 - 2x - 2tx + x^2 - 2tx^2 + t^2x^2}}{2} \\
&v \text{ is undefined} \\
w &= \frac{1 - \sqrt{1-4x}}{2}.
\end{aligned}
$$

From the simple form of w when $S = \mathbb{P}$ we deduce that $T_{\mathbb{P}}(n) = C_{n-1}$, another occurrence of Catalan numbers!

The problem of counting plane trees with no vertex of degree one by number of endpoints is equivalent to "Schröder's second problem" and is solved by the generating function (6.22). See the Notes for a surprising reference to this problem going back to the second century B.C. Now is a good time to give a general overview of Schröder's famous "*vier combinatorische Probleme*" (four combinatorial problems), since we have already given solutions to all four problems in Chapter 5 and the present chapter. Schröder was concerned with the enumeration of *bracketings*. He considered two classes of bracketings, viz., bracketings of words (or strings) and of sets, and two rules of combination (binary and arbitrary), giving four problems in all. A bracketing of a word $w = w_1w_2\cdots w_n$ is obtained by expressing w as a product of at least two nonempty words (unless w is a single letter), say $w = u_1u_2\cdots u_k$, and then inductively bracketing each u_i, continuing until only singleton words (letters) remain. Similarly, a bracketing of an n-set S is obtained by partitioning S into at least two (unless $\#S = 1$) nonempty pairwise disjoint subsets $S = T_1 \cup T_2 \cup \cdots \cup T_k$, and then inductively bracketing each T_i, continuing until only singletons remain. (The order of the T_i's is irrelevant.) A bracketing is *binary* if at each stage a non-singleton word or set is divided into exactly two parts; it is *arbitrary* if any number of parts (at least two) is allowed. Thus the four problems of Schröder are the following.

First Problem. Binary word bracketings. Clearly this coincides with binary parenthesizations of a string of length n, and was solved by Corollary 6.2.3(iii).

The generating function is given by

$$\sum_{n\geq 1} s_1(n)x^n = \frac{1-\sqrt{1-4x}}{2}$$
$$= x + x^2 + 2x^3 + 5x^4 + 14x^5 + 42x^6 + 132x^7 + 429x^8 + \cdots,$$

and

$$s_1(n) = C_{n-1} = \frac{1}{n}\binom{2(n-1)}{n-1},$$

the $(n-1)$-st Catalan number. For a generalization, see Exercise 5.43.

Second Problem. Arbitrary word bracketings. This is equivalent to plane trees with no vertex of degree one and n endpoints, with generating function (6.22) given by

$$\sum_{n\geq 1} s_2(n)x^n = \frac{1+x-\sqrt{1-6x+x^2}}{4}$$
$$= x + x^2 + 3x^3 + 11x^4 + 45x^5 + 197x^6 + 903x^7 + 4279x^8 + \cdots.$$

The numbers $r_n = 2s_2(n+1)$, $n \geq 1$ (with $r_0 = 1$), and $s_n = s_2(n+1)$, $n \geq 0$, are called *Schröder numbers.* Sometimes s_n is called a *little Schröder number* to distinguish it from r_n. Note that Proposition 6.2.1 yields several additional interpretations of Schröder numbers, e.g., s_n is the number of ways to dissect a convex $(n+2)$-gon with any number of diagonals that don't intesect in their interiors. For further information about Schröder numbers, see Exercise 6.39.

Third Problem. Binary set bracketings. This was the problem considered in Example 5.2.6 with exponential generating function

$$\sum_{n\geq 1} s_3(n)\frac{x^n}{n!} = 1-\sqrt{1-2x}$$
$$= x + \frac{x^2}{2!} + 3\frac{x^3}{3!} + 15\frac{x^4}{4!} + 105\frac{x^5}{5!} + 945\frac{x^6}{6!} + 10395\frac{x^7}{7!} + \cdots,$$

and $s_3(n) = 1\cdot 3\cdot 5\cdots(2n-3)$.

Fourth Problem. Arbitrary set bracketings. This problem was considered in Example 5.2.5, with exponential generating function

$$\sum_{n\geq 1} s_4(n)\frac{x^n}{n!} = (1+2x-e^x)^{\langle -1\rangle}$$
$$= x + \frac{x^2}{2!} + 4\frac{x^3}{3!} + 26\frac{x^4}{4!} + 236\frac{x^5}{5!} + 2752\frac{x^6}{6!}$$
$$+ 39208\frac{x^7}{7!} + 660032\frac{x^8}{8!} + \cdots.$$

See also Exercises 5.26 and 5.40.

6.3 Diagonals

There is a useful general method for obtaining algebraic generating functions that includes Example 6.2.7 and related results.

6.3.1 Definition. Let

$$F(x_1, \ldots, x_k) = \sum f(n_1, \ldots, n_k) x_1^{n_1} \cdots x_k^{n_k} \in K[[x_1, \ldots, x_k]].$$

The *diagonal* of F, denoted $\operatorname{diag} F$, is the power series in a single variable x defined by

$$\operatorname{diag} F = (\operatorname{diag} F)(x) = \sum_n f(n, n, \ldots, n) x^n.$$

6.3.2 Example. Write s and t for x_1 and x_2, and let

$$F(s,t) = \frac{1}{1-s-t} = \sum_{i,j} \binom{i+j}{i} s^i t^j.$$

Then

$$\operatorname{diag} F = \sum_n \binom{2n}{n} x^n = \frac{1}{\sqrt{1-4x}}.$$

Example 6.3.2 is a prototype for the following general result. It will be our main instance of the use of Puiseux's theorem.

6.3.3 Theorem. *Suppose $F(s,t) \in K[[s,t]] \cap K(s,t)$, i.e., F is a power series in s and t that represents a rational function. Then* $\operatorname{diag} F$ *is algebraic.*

Proof. We may assume K is algebraically closed (since F is algebraic over K if and only if F is algebraic over any algebraic extension of K). We also assume that $\operatorname{char} K = 0$, though the theorem in fact holds for any K. (See Exercise 6.11(b) for a more general result when $\operatorname{char} K = p > 0$.) Let

$$G = G(x,s) = F(s, x/s) \in K[[s, x/s]]. \tag{6.23}$$

Thus G is a formal Laurent series in s and x, such that if a monomial $x^i s^j$ appears in G, then $i \geq 0$ and $j \geq -i$. Note that

$$\operatorname{diag} F = [s^0]G,$$

the constant term of G regarded as a Laurent series in s whose coefficients are power series in x.

Since F is rational, it follows that $G = P/Q$, where $P, Q \in K[s, x]$. Regarding G as a rational function of s whose coefficients lie in $K[x]$ (or in the field $K(x)$),

we have a partial-fraction decomposition

$$G = \sum_{j=1}^{l} \frac{N_j(s)}{(s-\xi_j)^{e_j}}, \tag{6.24}$$

where (i) $e_j \in \mathbb{P}$, (ii) $\xi_1, \ldots, \xi_l$ are the distinct zeros of $Q(s)$, and (iii) $N_j(s) \in K(\xi_1, \ldots, \xi_l)[s]$. Since the coefficients of $Q(s)$ are polynomials in x, the zeros $\xi_1, \ldots, \xi_l$ are algebraic functions of x and hence by Puiseux's theorem (Theorem 6.1.5) we may assume $\xi_j \in K((x^{1/r}))$ for some $r \in \mathbb{P}$. Hence $N_j(s) \in K((x^{1/r}))[s]$.

Rename $\xi_1, \ldots, \xi_l$ as $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n$ so that

$$\alpha_i \in x^{1/r} K\big[\big[x^{1/r}\big]\big], \qquad \beta_j \notin x^{1/r} K\big[\big[x^{1/r}\big]\big].$$

In other words, the Puiseux expansion of α_i has only positive exponents, while that of β_j has some nonpositive exponent. Equivalently, $\beta_j^{-1} \in K[[x^{1/r}]]$. Thus

$$\begin{aligned} G &= \sum_{i=1}^{m} \frac{p_i(s)}{(1-s^{-1}\alpha_i)^{c_i}} + \sum_{j=1}^{n} \frac{q_j(s)}{\left(1-s\beta_j^{-1}\right)^{d_j}} \\ &= \sum_{i=1}^{m} p_i(s) \sum_{k\geq 0} \binom{c_i}{k} s^{-k}\alpha_i^k + \sum_{j=1}^{n} q_j(s) \sum_{k\geq 0} \binom{d_j}{k} s^k \beta_j^{-k} \end{aligned} \tag{6.25}$$

for some $c_i, d_j \in \mathbb{P}$ and $p_i(s), q_j(s) \in K((x^{1/r}))[s, s^{-1}]$. The expansion (6.25) for G satisfies

$$G \in K\big[\big[s^{1/r}, (x/s)^{1/r}\big]\big]\big[s^{-1/r}\big] \supset K[[s, x/s]].$$

(This is the key point – we must expand each term of (6.24) so that each expansion lies in a ring R (independent of the term) that contains $K[[s, x/s]]$, the ring in which the expansion (6.23) of G lies. Thus all our operations are taking place inside a single ring and hence are well defined.) Since $G \in K[[s, x/s]]$, the expansion (6.25) agrees with the series $G = F(s, x/s)$.

Now take the coefficient of s^0 on both sides of (6.25). The left-hand side becomes $\operatorname{diag} F$. The right-hand side becomes a finite sum of terms $\gamma\alpha^a$ and $\delta\beta^b$, where $a, b \in \mathbb{Z}$ and $\alpha, \beta, \gamma, \delta$ are algebraic. Hence $\operatorname{diag} F$ is algebraic, as was to be proved. □

6.3.4 Example. Consider the rational function $F(s,t) = (1-s-t)^{-1}$ of Example 6.3.2. Going through the previous proof yields

$$G = \frac{1}{1-s-\frac{x}{s}} = -\frac{s}{s^2-s+x}.$$

Now

$$s^2 - s + x = (s-\alpha)(s-\beta),$$

where

$$\alpha = \frac{1-\sqrt{1-4x}}{2} = x + \cdots$$

$$\beta = \frac{1+\sqrt{1-4x}}{2} = 1 + \cdots.$$

The partial-fraction expansion of G is

$$G = \frac{1}{\beta-\alpha}\left[\frac{\alpha}{s-\alpha} - \frac{\beta}{s-\beta}\right].$$

We must write this expression in a form so that its implied expansion as a fractional Laurent series (here just a Laurent series) agrees with the expansion

$$\begin{aligned} G = F(s, x/s) &= \sum_{i,j} \binom{i+j}{i} s^i \left(\frac{x}{s}\right)^j \\ &= \sum_{i,j} \binom{i+j}{i} x^j s^{i-j}. \end{aligned}$$

The difficulty is that a rational function such as $\alpha/(s-\alpha)$ has two expansions in powers of s, viz.,

$$\frac{\alpha/s}{1-\frac{\alpha}{s}} = \sum_{n\geq 1} \alpha^n s^{-n}$$

$$-\frac{1}{1-\frac{s}{\alpha}} = -\sum_{n\geq 0} \alpha^{-n} s^n.$$

The "correct" expansion will be the one that lies in the ring $K[[s^{1/r}, (x/s)^{1/r}]][s^{-1/r}]$ (here $r = 1$), viz., the first of the two possible expansions above (since $\alpha = x +$ higher order terms). Similarly we must regard $\beta/(s-\beta)$ as $-1/(1-s\beta^{-1})$, so (using $\beta - \alpha = \sqrt{1-4x}$)

$$\begin{aligned} G &= \frac{1}{\sqrt{1-4x}}\left(\frac{\alpha/s}{1-s^{-1}\alpha} + \frac{1}{1-s\beta^{-1}}\right) \\ &= \frac{1}{\sqrt{1-4x}}\left(\sum_{n\geq 1} \alpha^n s^{-n} + \sum_{n\geq 0} \beta^{-n} s^n\right). \end{aligned}$$

We read off immediately that

$$\operatorname{diag} F = [s^0]G = \frac{1}{\sqrt{1-4x}}.$$

It is natural to ask whether Theorem 6.3.3 has a proof avoiding fractional series. There is an elegant proof based on contour integration (really a variant of the proof we have just given) that we sketch for the benefit of readers knowledgeable about complex analysis. (For another proof, see the end of Section 6.7.) The contour integration approach has the advantage, as exemplified in Example 6.3.5 below, that one can use the vast arsenal of residue calculus (in particular, techniques for computing residues) to carry out the necessary computations. We could develop an equivalent theory for dealing with partial fractions and fractional series arising in our first proof of Theorem 6.3.3, but for those who know residue calculus such a development is unnecessary.

We assume $K = \mathbb{C}$. The series expansion of $F(s,t) = \sum f(m,n)s^m t^n$ converges for sufficiently small $|s|$ and $|t|$ (since $F(s,t)$ is rational and we are assuming it has a power series expansion about $s = t = 0$). Thus $\operatorname{diag} F$ will converge for $|x|$ small. Fix such a small x. Then the series

$$F(s, x/s) = \sum f(m,n)s^{m-n}x^n,$$

regarded as a function of s, will converge in an annulus about $s = 0$ and hence on some circle $|s| = \rho > 0$. By Cauchy's integral theorem,

$$\operatorname{diag} F = [s^0]F(s, x/s) = \frac{1}{2\pi i}\int_{|s|=\rho} F(s, x/s)\frac{ds}{s}.$$

By the Residue Theorem,

$$\operatorname{diag} F = \sum_{s=s(x)} \operatorname{Res}_s F(s, x/s)\frac{1}{s},$$

where the sum ranges over all singularities of $\frac{1}{s}F(s, x/s)$ inside the circle $|s|=\rho$. (Such singularities are precisely the ones satisfying $\lim_{x\to 0} s(x) = 0$.) Since $F(s, x/s)$ is rational, all such singularities $s(x)$ are poles, and $s(x)$ is an algebraic function of x. The residue at a pole s belongs to the ring $\mathbb{C}(s, x)$, so the residues are algebraic. Hence $\operatorname{diag} F$ is algebraic.

6.3.5 Example. Let us apply the above proof to our canonical example $F(s,t) = (1-s-t)^{-1}$. We have

$$\operatorname{diag} F = \frac{1}{2\pi i}\int \frac{ds}{s\left(1-s-\frac{x}{s}\right)} = \frac{1}{2\pi i}\int \frac{ds}{-x+s-s^2}.$$

The poles are at $s = \frac{1}{2}(1 \pm \sqrt{1-4x})$. The only pole approaching 0 as $x \to 0$ is $s_0 = \frac{1}{2}(1 - \sqrt{1-4x})$. If $A(s)/B(s)$ has a simple pole at s_0 and $A(s_0) \neq 0$, then the residue at s_0 is $A(s_0)/B'(s_0)$. Hence

$$\operatorname{diag} F = \operatorname{Res}_{s_0} \frac{1}{-x+s-s^2} = \frac{1}{1-2s_0} = \frac{1}{\sqrt{1-4x}}.$$

6.3.6 Example. Let us consider a somewhat more complicated example than the preceding, where it is not feasible to find the poles explicitly. Consider the series $z = \sum_{n\geq 0} \binom{kn}{n} x^n$ of Example 6.2.7, where $k \in \mathbb{P}$. The contour integration argument given above for computing diagonals easily extends to give

$$\begin{aligned} z &= \frac{1}{2\pi i} \int_{s=|\rho|} \frac{ds}{s\left(1 - \frac{u}{s} - s^{k-1}\right)} \\ &= \frac{1}{2\pi i} \int_{s=|\rho|} \frac{ds}{-u+s-s^k} \end{aligned} \qquad (6.26)$$

for a suitable $\rho > 0$, where $u = \sqrt[k-1]{x}$. Since the polynomial $-u+s-s^k$ has a zero of multiplicity one at $s=0$ when we set $x=0$, it follows that the integrand of (6.26) has a single (simple) pole $s_0 = s_0(x)$ satisfying $\lim_{x\to 0} s_0(x) = 0$. Thus

$$\begin{aligned} z &= \operatorname{Res}_{s=s_0} \frac{1}{-u+s-s^k} \\ &= \frac{1}{1-ks_0^{k-1}}. \end{aligned}$$

Hence $1/z = 1 - ks_0^{k-1}$, so $s_0/z = s_0 - ks_0^k = s_0 - k(s_0 - u)$. Solving for s_0 gives

$$s_0 = \frac{kuz}{1+(k-1)z},$$

so from $s_0^k - s_0 + u = 0$ we get

$$\left(\frac{kuz}{1+(k-1)z}\right)^k - \frac{kuz}{1+(k-1)z} + u = 0.$$

Dividing by u gives

$$x\left(\frac{kz}{1+(k-1)z}\right)^k - \frac{kz}{1+(k-1)z} + 1 = 0.$$

This equation is equivalent to (6.21) and simplifies to a polynomial equation satisfied by z.

In view of Theorem 6.3.3, it's natural to ask whether diag F is algebraic for a rational series in more than two variables. Unfortunately the answer in general is

negative; we have mentioned earlier that $u = \sum \binom{3n}{n,n,n} x^n$ is not algebraic, yet $u = \operatorname{diag}(1 - x_1 - x_2 - x_3)^{-1}$. But diagonals of rational series in any number of variables do possess the desirable property of D-finiteness, to be discussed in the next section; see Exercise 6.61 for the connection between diagonals and D-finiteness.

We conclude this section with a class of enumerative problems that involve in a natural way rational functions of two variables. The general setting is that of *lattice paths in the plane*. For simplicity we will deal with paths whose steps come from $\mathbb{N} \times \mathbb{N}$, though many variations of this condition are possible. Given $S \subseteq \mathbb{N} \times \mathbb{N}$, an *$S$-path* of length l from $(0, 0)$ to (m, n) is a sequence $\sigma = (v_1, v_2, \ldots, v_l) \in S^l$ such that $v_1 + v_2 + \cdots + v_l = (m, n)$. Thus σ may be regarded as a *weak* (since $(0, 0) \in S$ is allowed) *composition* (or ordered partition) of (m, n) into l parts – the exact two-dimensional analogue of the one-dimensional situation discussed in Section 1.2. The set S is the set of *allowed steps*. Let $N_S(m, n; l)$ denote the number of S-paths of length l from $(0, 0)$ to (m, n). If $(0, 0) \notin S$ then let

$$N_S(m, n) = \sum_l N_S(m, n; l),$$

the total number of S-paths from $(0, 0)$ to (m, n). Define the generating function

$$F_S(s, t; z) = \sum_{m,n,l} N_S(m, n; l) s^m t^n z^l,$$

and if $(0, 0) \notin S$ then define

$$G_S(s, t) = F_S(s, t; 1) = \sum_{m,n} N_S(m, n) s^m t^n.$$

6.3.7 Proposition. *Let $S \subseteq \mathbb{N} \times \mathbb{N}$. Then*

$$F_S(s, t; z) = \frac{1}{1 - z \sum_{(i,j) \in S} s^i t^j}.$$

If moreover $(0, 0) \notin S$, *then*

$$G_S(s, t) = \frac{1}{1 - \sum_{(i,j) \in S} s^i t^j}.$$

Proof. It should be clear that for fixed l,

$$\sum_{m,n} N_S(m, n; l) s^m t^n = \left(\sum_{(i,j) \in S} s^i t^j \right)^l.$$

Now multiply by z^l and sum on $l \geq 0$. □

Under certain circumstances F_S and G_S will be rational functions of s and t, and their diagonals (or diagonals of closely related series) will be algebraic. Let us give some examples.

6.3.8 Example. Exercise 1.5(b) asked for the generating function $y = \sum f(n)x^n$, where $f(n)$ counts the number of lattice paths from (0, 0) to (n, n) with steps (1, 0), (0, 1), and (1, 1). The solution $y = 1/\sqrt{1-6x+x^2}$ was obtained by an *ad hoc* argument, and the final answer appears rather mysterious. Using Proposition 6.3.7 and our techniques for computing diagonals, we have

$$\begin{aligned} y &= \operatorname{diag} \frac{1}{1-s-t-st} \\ &= [s^0]\frac{1}{1-s-\frac{x}{s}-s\frac{x}{s}} \\ &= [s^0]\frac{-s}{s^2+(x-1)s+x} \\ &= [s^0]\frac{1}{\beta-\alpha}\left(\frac{s}{s-\alpha}-\frac{s}{s-\beta}\right), \end{aligned}$$

where $\alpha = \frac{1}{2}(1-x-\sqrt{1-6x+x^2})$ and $\beta = \frac{1}{2}(1-x+\sqrt{1-6x+x^2})$. Thus

$$\begin{aligned} y &= [s^0]\frac{1}{\sqrt{1-6x+x^2}}\left(\frac{s\alpha^{-1}}{1-s\alpha^{-1}}+\frac{1}{1-s^{-1}\beta}\right) \\ &= \frac{1}{\sqrt{1-6x+x^2}} \\ &= 1+3x+13x^2+63x^3+321x^4+1683x^5+8989x^6+\cdots. \end{aligned} \tag{6.27}$$

The coefficient of $s^m t^n$ in the generating function $1/(1-s-t-st)$ is known as the *Delannoy number* $D(m,n)$, and $D(n,n)$ is called a *central Delannoy number*. Thus in particular

$$\sum_{n\geq 0} D(n,n)x^n = \frac{1}{\sqrt{1-6x+x^2}}.$$

We could also take into account the number of steps, i.e., now let

$$y = \sum_{n,l} N_S(n,n;l)x^n z^l,$$

where $S = \{(1,0),(0,1),(1,1)\}$. Then one can compute that

$$\begin{aligned} y &= \operatorname{diag} \frac{1}{1-z(s+t+st)} \\ &= \frac{1}{\sqrt{1-x(2z+4z^2)+x^2z^2}}, \end{aligned} \tag{6.28}$$

where diag is computed with respect to the variables s and t only (not z). In

particular, if $f(l)$ is the number of paths that end up on the line $y = x$ after l steps, then setting $x = 1$ in (6.28) yields

$$\begin{aligned}\sum_{l\geq 0} f(l)z^l &= \frac{1}{\sqrt{1-2z-3z^2}}\\ &= 1+z+3z^2+7z^3+19z^4+51z^5+141z^6\\ &\quad +393z^7+1107z^8+3139z^9+8953z^{10}+\cdots. \qquad (6.29)\end{aligned}$$

For some more information on this generating function, see Exercise 6.42.

If we were even more ambitious, we could keep track of the number of steps of each type (i.e., $(1,0)$, $(0,1)$, $(1,1)$). Letting z_{ij} keep track of the number of times (i,j) is a step, it is clear that we want to compute

$$\begin{aligned}y &= \operatorname{diag}_{s,t} \frac{1}{1-z_{10}s-z_{01}t-z_{11}st}\\ &= \frac{1}{\sqrt{1-x(2z_{11}+4z_{10}z_{01})+x^2z_{11}^2}},\end{aligned}$$

where $\operatorname{diag}_{s,t}$ denotes the diagonal with respect to the variables s and t only. In this case we gain no further information than contained in (6.28), since knowing the destination (n,n) and the total number of steps determines the number of steps of each type, but for less "homogeneous" choices of S this is no longer the case.

6.3.9 Example. Set $S = \mathbb{N}\times\mathbb{N} - \{(0,0)\}$, so $N_S(n,n)$ is the total number of paths from $(0,0)$ to (n,n) with any allowed lattice steps that move closer to (n,n) (and never pass (n,n) and then backtrack). We have

$$\begin{aligned}y := \sum_{n\geq 0} N_S(n,n)x^n &= \operatorname{diag} \frac{1}{1-\left(\frac{1}{(1-s)(1-t)}-1\right)}\\ &= \operatorname{diag} \frac{(1-s)(1-t)}{1-2s-2t+2st}.\end{aligned}$$

By now it should be routine to compute that

$$\begin{aligned}y &= \frac{1}{2}\left(1+\frac{1}{\sqrt{1-12x+4x^2}}\right) \qquad (6.30)\\ &= 1+3x+26x^2+252x^3+2568x^4+26928x^5+287648x^6+\cdots.\end{aligned}$$

(The reader may find it instructive to carry out the computation in order to ascertain how the term $\frac{1}{2}$ arises in (6.30) when y is written as $\frac{1}{2}+\frac{1}{2\sqrt{1-12x+4x^2}}$.)

6.4 D-Finite Generating Functions

An important property of rational generating functions $\sum f(n)x^n$ is that the coefficients $f(n)$ satisfy a simple recurrence relation (Theorem 4.1.1(ii)). An analogous result holds for algebraic generating functions. However, the type of recurrence satisfied by the coefficients of an algebraic generating function also holds for the coefficients of more general series, whose basic properties we now discuss.

6.4.1 Proposition. *Let $u \in K[[x]]$. The following three conditions are equivalent:*

(i) *The vector space over $K(x)$ spanned by u and all its derivatives $u', u'', \ldots$ is finite-dimensional. In symbols,*

$$\dim_{K(x)}[K(x)u + K(x)u' + K(x)u'' + \cdots] < \infty.$$

(ii) *There exist polynomials $p_0(x), \ldots, p_d(x) \in K[x]$ with $p_d(x) \neq 0$, such that*

$$p_d(x)u^{(d)} + p_{d-1}(x)u^{(d-1)} + \cdots + p_1(x)u' + p_0(x)u = 0, \quad (6.31)$$

where $u^{(j)} = d^j u/dx^j$.

(iii) *There exist polynomials $q_0(x), \ldots, q_m(x), q(x) \in K[x]$, with $q_m(x) \neq 0$, such that*

$$q_m(x)u^{(m)} + q_{m-1}(x)u^{(m-1)} + \cdots + q_1(x)u' + q_0(x)u = q(x). \quad (6.32)$$

*If u satisfies any (and hence all) of the above three conditions, then we say that u is a D-*finite *(short for* differentiably finite*) power series.*

Proof. (i) $\Rightarrow$ (ii). Suppose

$$\dim_{K(x)}[K(x)u + K(x)u' + \cdots] = d.$$

Thus $u, u', \ldots, u^{(d)}$ are linearly dependent over $K(x)$. Write down a dependence relation and clear denominators so that the coefficients are all polynomials to get an equation of the form (6.31).

(ii) $\Rightarrow$ (iii). Trivial.

(iii) $\Rightarrow$ (i). Suppose that (6.32) holds and that the usual degree of $q(x)$ as a polynomial in x is t. Differentiate (6.32) $t+1$ times to get an equation (6.31), with $d = m+t+1$ and $p_d(x) = q_m(x) \neq 0$. Solving for $u^{(d)}$ shows that

$$u^{(d)} \in K(x)u + K(x)u' + \cdots + K(x)u^{(d-1)}. \quad (6.33)$$

Differentiate (with respect to x) the equation expressing $u^{(d)}$ as a $K(x)$-linear

combination of $u, u', \ldots, u^{(d-1)}$. We get

$$\begin{aligned} u^{(d+1)} &\in K(x)u + K(x)u' + \cdots + K(x)u^{(d)} \\ &= K(x)u + K(x)u' + \cdots + K(x)u^{(d-1)} \qquad \text{(by (6.33))}. \end{aligned}$$

Continuing in this way, we get

$$u^{(d+k)} \in K(x)u + K(x)u' + \cdots + K(x)u^{(d-1)}$$

for all $k \geq 0$, so (i) holds. □

6.4.2 Example. (a) $u = e^x$ is D-finite, since $u' = u$. Similarly, any linear combination of series of the form $x^m e^{ax}$ ($m \in \mathbb{N}$, $a \in K$) is D-finite, since such series satisfy a linear homogeneous differential equation with *constant* coefficients.
(b) $u = \sum_{n\geq 0} n! x^n$ is D-finite, since $(xu)' = \sum_{n\geq 0}(n+1)! x^n$, whence $1 + x(xu)' = u$. This equation simplifies to $x^2 u' + (x-1)u = -1$.

We wish to characterize the coefficients of a D-finite power series u. To this end, define a function $f : \mathbb{N} \to K$ to be *P-recursive* (short for *polynomially recursive*) if there exist polynomials $P_0, \ldots, P_e \in K[n]$ with $P_e \neq 0$, such that

$$P_e(n)f(n+e) + P_{e-1}(n)f(n+e-1) + \cdots + P_0(n)f(n) = 0, \qquad (6.34)$$

for all $n \in \mathbb{N}$. In other words, f satisfies a homogeneous linear recurrence of finite degree with polynomial coefficients.

6.4.3 Proposition. *Let $u = \sum_{n\geq 0} f(n)x^n \in K[[x]]$. Then u is D-finite if and only if f is P-recursive.*

Proof. Suppose u is D-finite, so that (6.31) holds (with $p_d(x) \neq 0$). Since

$$x^j u^{(i)} = \sum_{n\geq 0} (n+i-j)_i f(n+i-j)x^n,$$

when we equate coefficients of x^{n+k} in (6.31) for fixed k sufficiently large, we will obtain a recurrence of the form (6.34) for f. This recurrence will not collapse to $0 = 0$, because if $[x^j]p_d(x) \neq 0$, then $[n^d]P_{d-j+k}(n) \neq 0$.

Conversely, suppose that f satisfies (6.34) (with $P_e(n) \neq 0$). For fixed $i \in \mathbb{N}$, the polynomials $(n+i)_j$, $j \geq 0$, form a K-basis for the space $K[n]$ (since $\deg(n+i)_j = j$). Thus $P_i(n)$ is a K-linear combination of the polynomials $(n+i)_j$, so $\sum_{n\geq 0} P_i(n)f(n+i)x^n$ is a K-linear combination of series of the

form $\sum_{n\geq 0}(n+i)_j f(n+i)x^n$. Now

$$\sum_{n\geq 0}(n+i)_j f(n+i)x^n = R_i(x) + x^{j-i}u^{(j)}$$

for some $R_i(x) \in x^{-1}K[x^{-1}]$ (i.e., $R_i(x)$ is a Laurent polynomial all of whose exponents are *negative*). For instance, $u' = \sum_{n\geq 0} nf(n)x^{n-1}$, so

$$\begin{aligned} x^{-1}u' &= \sum_{n\geq -1}(n+2)f(n+2)x^n \\ &= f(1)x^{-1} + \sum_{n\geq 0}(n+2)f(n+2)x^n. \end{aligned}$$

Hence multiplying (6.34) by x^n and summing on $n \geq 0$ yields

$$0 = \sum a_{ij}x^{j-i}u^{(j)} + R(x), \tag{6.35}$$

where the sum is finite, $a_{ij} \in K$, and $R(x) \in x^{-1}K[x^{-1}]$. One easily sees that not all $a_{ij} = 0$. Now multiply (6.35) by x^q for q sufficiently large to get an equation of the form (6.32). □

6.4.4 Example. (a) $f(n) = n!$ is clearly P-recursive, since $f(n+1) - (n+1)f(n) = 0$. Hence $u = \sum_{n\geq 0} n!x^n$ is D-finite. We don't need a trick to see this as in Example 6.4.2(b).
(b) $f(n) = \binom{2n}{n}$ is P-recursive, since

$$(n+1)f(n+1) - 2(2n+1)f(n) = 0.$$

Hence $u = \sum_{n\geq 0}\binom{2n}{n}x^n = 1/\sqrt{1-4x}$ is D-finite.

Soon we will give many examples of D-finite series and P-recursive functions (Theorems 6.4.6, 6.4.9, 6.4.10, 6.4.12 and Exercises 6.53–6.61). First we note a simple but useful result.

6.4.5 Proposition. *Suppose $f : \mathbb{N} \to K$ is P-recursive, and $g : \mathbb{N} \to K$ agrees with f for all n sufficiently large. Then g is P-recursive.*

Proof. Suppose $f(n) = g(n)$ for $n \geq n_0$ and f satisfies (6.34). Then

$$\left(\prod_{j=0}^{n_0-1}(n-j)\right)[P_e(n)g(n+e) + \cdots + P_0(n)g(n)] = 0,$$

so g is P-recursive. Alternatively, we could use

$$P_e(n+n_0)g(n+e+n_0) + P_{e-1}(n+n_0)g(n+e+n_0-1)$$
$$+\cdots+P_0(n+n_0)g(n+n_0) = 0. \qquad \square$$

Our first main result on D-finite power series asserts that algebraic series are D-finite.

6.4.6 Theorem. *Let $u \in K[[x]]$ be algebraic of degree d. Then u is D-finite. More precisely, u satisfies an equation (6.31) of order d, or an equation (6.32) of order $m = d-1$. (For the least-order differential equation satisfied by u, see Exercise 6.62.)*

Proof. By (6.12) we have $u' \in K(x,u)$. Continually differentiating (6.12) with respect to x shows by induction that $u^{(k)} \in K(x,u)$ for all $k \geq 0$. But $\dim_{K(x)} K(x,u) = d$, so $u, u', \ldots, u^{(d)}$ are linearly dependent over $K(x)$, yielding an equation of the form (6.31). Similarly, $1, u, u', \ldots, u^{(d-1)}$ are linearly dependent over $K(x)$, yielding an equation (6.32) with $m \leq d-1$. $\square$

Proposition 6.4.3 and Theorem 6.4.6 together show that the coefficients of an algebraic power series u satisfy a simple recurrence (6.34). In particular, once this recurrence is found (which involves only a finite amount of computation), we have a method for rapidly computing the coefficients of u. Let us note that not all D-finite series are algebraic, e.g., e^x (Exercise 6.1).

6.4.7 Example. Let $F(x) \in K(x)$ with $\operatorname{char} K = 0$ and $F(0) = 1$ (so $F(x)^{1/d}$ is defined formally as a power series for any $d \in \mathbb{P}$). Let $u = F(x)^{1/d}$ (with $u(0) = 1$), so $u^d = F(x)$. Then $du^{d-1}u' = F'(x)$, so multiplying by u yields

$$dF(x)u' = F'(x)u. \tag{6.36}$$

If we want polynomial coefficients in (6.36), suppose $F(x) = A(x)/B(x)$ where $A, B \in K[x]$. Then (6.36) becomes

$$dABu' = (A'B - AB')u. \tag{6.37}$$

For instance, if $B(x) = 1$, so $F(x) = A(x) = a_0 + a_1x + \cdots + a_rx^r$; and if $u = \sum f(n)x^n$, then equating coefficients of x^{n+r-1} in (6.37) yields

$$\sum_{j=0}^{r} a_{r-j}(dn + dj - r + j)f(n+j) = 0. \tag{6.38}$$

Similarly, if $A(x) = 1$, $B(x) = b_0 + b_1x + \cdots + b_rx^r$, and $u = \sum f(n)x^n$, then

$$\sum_{j=0}^{r} b_{r-j}(dn + dj + r - j)f(n+j) = 0. \tag{6.39}$$

6.4.8 Example. (a) In Exercise 1.5(b) and Example 6.3.8 we saw that the number $f(n)$ of lattice paths from $(0, 0)$ to (n, n) with steps $(1, 0)$, $(0, 1)$, and $(1, 1)$ satisfied

$$u := \sum_{n\geq 0} f(n)x^n = \frac{1}{\sqrt{1-6x+x^2}}.$$

Putting $B(x) = 1 - 6x + x^2$ and $d = 2$ in (6.39) yields

$$(n+2)f(n+2) - 3(2n+3)f(n+1) + (n+1)f(n) = 0 \qquad (6.40)$$

for $n \geq 0$, with the initial conditions $f(0) = 1$, $f(1) = 3$. Though equation (6.40) is a simple recurrence, it is difficult to give a combinatorial proof.
(b) Here are the recurrences satisfied by the coefficients $f(n)$ of some of the algebraic functions $u = \sum f(n)x^n$ considered earlier in the chapter:

(i) $u = \frac{1+x-\sqrt{1-6x+x^2}}{4}$ (equation (6.22)):

$$(n+2)f(n+2) - 3(2n+1)f(n+1) + (n-1)f(n) = 0, \qquad n \geq 1.$$

(ii) $u = \frac{1}{\sqrt{1-2x-3x^2}}$ (equation (6.29)):

$$(n+2)f(n+2) - (2n+3)f(n+1) - 3(n+1)f(n) = 0, \qquad n \geq 0.$$

(iii) $u = \frac{1}{2}\left(1 + \frac{1}{\sqrt{1-12x+4x^2}}\right)$ (equation (6.30)):

$$(n+2)f(n+2) - 6(2n+3)f(n+1) + 4(n+1)f(n) = 0, \qquad n \geq 1.$$

Theorem 6.4.6 yields a large class of interesting D-finite series. But there are many simple series for which it is unclear, based on what we have presented up to now, whether they are D-finite, such as

$$\begin{aligned}
u_1 &= \sec x \\
u_2 &= \sqrt{\cos x} \\
u_3 &= e^{x/\sqrt{1-4x}} \\
u_4 &= e^{e^x-1} \\
u_5 &= \frac{e^x}{\sqrt{1-4x}} \\
u_6 &= e^x + \sum_{n\geq 0} n!x^n \\
u_7 &= \sum_{n\geq 0}\left[1 + \binom{n}{1}^3 + \binom{n}{2}^3 + \cdots + \binom{n}{n}^3\right]^2 x^n.
\end{aligned}$$

Thus we need some further techniques for showing that power series are D-finite (or that their coefficients are P-recursive). It will follow from Theorem 6.4.10 that u_3 is D-finite, from Theorem 6.4.9 that u_5 and u_6 are D-finite, and from Theorem 6.4.12 that u_7 is D-finite. On the other hand, u_1, u_2, and u_4 are not D-finite. We will not develop here systematic methods for showing that power series are *not* D-finite, but see Exercises 6.59–6.60 for some tricks and results in that direction.

6.4.9 Theorem. *The set $\mathcal{D}$ of D-finite power series $u \in K[[x]]$ forms a subalgebra of $K[[x]]$. In other words, if $u, v \in \mathcal{D}$, and $\alpha, \beta \in K$, then $\alpha u + \beta v \in \mathcal{D}$ and $uv \in \mathcal{D}$.*

Proof. If $w \in K[[x]] \subset K((x))$, then let V_w denote the vector space over $K(x)$ spanned by $w, w', w'', \ldots$. Thus V_w is a subspace of $K((x))$. Let $u, v \in \mathcal{D}$ and $\alpha, \beta \in K$. Set $y = \alpha u + \beta v$. Then $y, y', y'', \ldots \in V_u + V_v$. Thus, taking dimensions over $K(x)$, we have

$$\dim V_y \leq \dim(V_u + V_v) \leq \dim V_u + \dim V_v < \infty.$$

Hence y is D-finite.

It remains to show that if $u, v \in \mathcal{D}$ then $uv \in \mathcal{D}$. We assume knowledge of the elementary properties of the tensor (or Kronecker) product of two vector spaces. Let $V = K((x))$, regarded as a vector space over $K(x)$. There is a unique linear transformation

$$\phi : V_u \otimes_{K(x)} V_v \to V$$

that satisfies $\phi(u^{(i)} \otimes v^{(j)}) = u^{(i)}v^{(j)}$ for all $i, j \geq 0$ (or $\phi(y \otimes z) = yz$ for all $y \in V_u, z \in V_v$). By Leibniz's rule for differentiating a product, we see that the image of ϕ contains V_{uv}. Hence

$$\dim V_{uv} \leq \dim(V_u \otimes V_v) = (\dim V_u)(\dim V_v) < \infty.$$

Thus $uv \in \mathcal{D}$, as desired. □

Our next result deals with the composition of D-finite series. In general, if $u, v \in \mathcal{D}$ with $u(0) = 0$, then $u(v(x))$ need not be D-finite, even if u is algebraic. For instance (Exercise 6.59) if $u = \sqrt{1+x} \in K_{\text{alg}}[[x]] \subset \mathcal{D}$ and $v = \log(1 + x^2) - 1 \in \mathcal{D}$, then $u(v(x)) \notin \mathcal{D}$. Thus the next result may come as something of a surprise.

6.4.10 Theorem. *If $u \in \mathcal{D}$ and $v \in K_{\text{alg}}[[x]]$ with $v(0) = 0$, then $u(v(x)) \in \mathcal{D}$.*

Proof. Let $y = u(v(x))$ and $i \geq 0$. By iterating the chain rule and Leibniz's rule, we see that $y^{(i)}$ is a linear combination of $u(v(x)), u'(v(x)), u''(v(x)), \ldots$ with coefficients in $K[v', v'', v''', \ldots]$. Since v is algebraic, the proof of Theorem 6.4.6 shows that $v^{(i)} \in K(x, v)$. Hence

$$K[v', v'', v''', \ldots] \subset K(x, v).$$

Let V be the $K(x, v)$-vector space spanned by $u(v(x)), u'(v(x)), \ldots$. Since $u \in \mathcal{D}$, we have

$$\dim_{K(x)} \operatorname{span}_{K(x)}\{u(x), u'(x), \ldots\} < \infty.$$

Thus

$$\dim_{K(v)} \operatorname{span}_{K(v)}\{u(v(x)), u'(v(x)), \ldots\} < \infty,$$

so *a fortiori*

$$\dim_{K(x,v)} \operatorname{span}_{K(v)}\{u(v(x)), u'(v(x)), \ldots\} < \infty.$$

The above argument shows that

$$\dim_{K(x,v)} V < \infty \quad \text{and} \quad [K(x, v) : K(x)] < \infty,$$

where $[L : M]$ denotes the degree of the field L over the subfield M (i.e., $[L : M] = \dim_M L$). Hence

$$\dim_{K(x)} V = \left(\dim_{K(x,v)} V\right) \cdot [K(x, v) : K(x)] < \infty.$$

Since each $y^{(i)} \in V$, there follows $y \in \mathcal{D}$. $\square$

6.4.11 Example. Since $\sum_{n\geq 0} n!x^n \in \mathcal{D}$ (by Example 6.4.2(b)), $e^x \in \mathcal{D}$ (by Example 6.4.2(a)), and $x/\sqrt{1-4x}$ is algebraic, we conclude from Theorems 6.4.9 and 6.4.10 that

$$u := \left(\sum_{n\geq 0} n!x^n\right) e^{x/\sqrt{1-4x}} \in \mathcal{D}.$$

It would be a chore, however, to find a differential equation (6.31) or (6.32) satisfied by u, or a linear recurrence (6.34) satisfied by its coefficients.

For our final basic result on D-finite series, we consider the Hadamard product. We will show that if $u, v \in \mathcal{D}$ then $u * v \in \mathcal{D}$. Equivalently, if $f(n)$ and $g(n)$ are P-recursive, then so is $f(n)g(n)$. The proof will be quite similar to the proof that products of D-finite series are D-finite (see Theorem 6.4.9). We would like to work with the set $\mathcal{P}$ of all P-recursive functions $f : \mathbb{N} \to K$, regarded as a vector space over the field $K(n)$ of rational functions in the variable n. In order for $\mathcal{P}$ to be a $K(n)$-vector space, we need to have that if $f, g \in \mathcal{P}$ and $R(n) = \frac{A(n)}{B(n)} \in K(n)$ (with $A, B \in K[n]$), then $f + g \in \mathcal{P}$ and $Rf \in \mathcal{P}$. It follows immediately from Theorem 6.4.9 that $f + g \in \mathcal{P}$. If, on the other hand, f satisfies (6.34), then

$h := Rf$ satisfies

$$\frac{P_e(n)B(n+e)h(n+e)}{A(n+e)} + \frac{P_{e-1}(n)B(n+e-1)h(n+e-1)}{A(n+e-1)}$$
$$+ \cdots + \frac{P_0(n)B(n)h(n)}{A(n)} = 0.$$

Multiplying by $A(n)A(n+1)\cdots A(n+e)$ yields a nonzero linear recurrence with polynomial coefficients satisfied by h, so $h \in \mathcal{P}$. There is one technical flaw in this argument, however. (Can the reader find it without reading further?) The problem is that $B(n)$ may have zeros at certain $n_0 \in \mathbb{N}$, so $h(n)$ is undefined at $n = n_0$. However, since $B(n)$ can have only finitely many zeros, $h(n)$ is defined for all sufficiently large n. Thus we must deal not with $h(n)$ itself, but only with its behavior "in a neighborhood of ∞." (An alternative approach is to regard $\mathcal{P}$ as a $K[n]$-module, rather than a $K(n)$-module. However, we have tried to keep algebraic prerequisites at a minimum by working with vector spaces as much as possible.)

To make the above ideas precise, define a relation $\sim$ on functions $h_1, h_2 : \mathbb{N} \to K$ by the rule that $h_1 \sim h_2$ if $h_1(n) = h_2(n)$ for all sufficiently large n. Clearly $\sim$ is an equivalence relation; we call the equivalence classes *germs* (more precisely, germs at ∞ of functions $h : \mathbb{N} \to K$). Denote the class containing h by $[h]$. We define linear combinations (over K) and multiplication of germs by $\alpha_1[h_1] + \alpha_2[h_2] = [\alpha_1 h_1 + \alpha_2 h_2]$, $[h_1][h_2] = [h_1 h_2]$. Clearly these operations are well defined. Given $R(n) \in K(n)$ and $h : \mathbb{N} \to K$, we can also define a germ $[Rh]$ by requiring that $[Rh] = [h_1]$, where $h_1 : \mathbb{N} \to K$ is any function agreeing with Rh for those n for which $R(n)$ is defined. In this way the space $\mathcal{G}$ of germs acquires the structure of a vector space over the field $K(n)$. Finally note that if $h_1 \sim h_2$, then $h_1 \in \mathcal{P}$ if and only if $h_2 \in \mathcal{P}$ by Proposition 6.4.5. Hence we may speak of *P-recursive germs*. Thus a germ $[h]$ is P-recursive if and only if the $K(n)$-vector subspace

$$\mathcal{G}_h = \mathrm{span}_{K(n)}\{[h(n)], [h(n+1)], [h(n+2)], \ldots\}$$

of $\mathcal{G}$ is finite-dimensional. We are now ready to state and prove the desired result.

6.4.12 Theorem. *If $f, g : \mathbb{N} \to K$ are P-recursive, then so is the product fg. Equivalently, if $u, v \in \mathcal{D}$ then $u * v \in \mathcal{D}$.*

Proof. The above discussion shows that it suffices to prove that if $[f]$ and $[g]$ are P-recursive germs, then so is $[fg]$. There is a unique linear transformation

$$\phi : \mathcal{G}_f \otimes_{K(n)} \mathcal{G}_g \to \mathcal{G}$$

that satisfies

$$\phi([f(n+i)] \otimes [g(n+j)]) = [f(n+i)][g(n+j)] = [f(n+i)g(n+j)].$$

Clearly the image of ϕ contains $\mathcal{G}_{fg} = \operatorname{span}_K\{[f(n)g(n)], [f(n+1)g(n+1)], \ldots\}$. Hence (taking dimensions over $K(n)$),

$$\dim \mathcal{G}_{fg} \leq \dim(\mathcal{G}_f \otimes \mathcal{G}_g) = (\dim \mathcal{G}_f)(\dim \mathcal{G}_g) < \infty,$$

so fg is P-recursive. □

6.5 Noncommutative Generating Functions

A powerful tool for showing that power series $\eta \in K[[x]]$ or $\eta \in K[[x_1, \ldots, x_m]]$ are rational or algebraic is the theory of *noncommutative* formal series (in several variables). The connections with rational power series is more or less equivalent to the transfer-matrix method (Section 4.7) and so won't yield any really new (commutative) rational generating functions. Similarly, it is possible to develop an analogue of the transfer-matrix method for (commutative) algebraic generating functions, so that noncommutative series are not really needed. However, the noncommutative approach to both rational and algebraic commutative generating functions yields an elegant and natural conceptual framework which can greatly simplify complicated computations. We will give an overview of both rational and algebraic noncommutative series. In one nice application (Theorem 6.7.1) we will use the theory of both rational and algebraic series. We will adhere to standard terminology and notation in this area, though it will be slightly different from our previous terminology and notation involving commutative series.

Let K denote a fixed field. (Much of the theory can be developed over an arbitrary "semiring" (essentially a ring without additive inverses, such as $\mathbb{N}$), and that generalization has some interesting features, but for our purposes a field will suffice.) Let X be a set, called an *alphabet*, and let X^* be the free monoid generated by X, as defined in Section 4.7. Thus X^* consists of all finite strings (including the empty string 1) $w_1 \cdots w_n$ of letters $w_i \in X$. We write $|w| = n$ if $w = w_1 \cdots w_n \in X^*$, with each $w_i \in X$. Also define $X^+ = X^* \backslash 1$ $(= X^* - \{1\})$.

6.5.1 Definition. A *formal* (*power*) *series* in X (over K) is a function $S : X^* \to K$. We write $\langle S, w\rangle$ for $S(w)$ and then write

$$S = \sum_{w \in X^*} \langle S, w\rangle w.$$

The set of all formal series in X is denoted $K\langle\langle X\rangle\rangle$.

The set $K\langle\langle X\rangle\rangle$ has the obvious structure of a ring (or even a K-algebra) with identity 1. (We identify $1 \in X^*$ with $1 \in K$, and abbreviate the term $\alpha \cdot 1$ of the

above series S as α.) Addition is componentwise, i.e.,

$$S + T = \sum_w (\langle S, w\rangle + \langle T, w\rangle) w,$$

while multiplication is given by the usual power series product, taking into account the noncommutativity of the variables. Thus

$$\begin{aligned}\left(\sum \langle S, u\rangle u\right)\left(\sum \langle T, v\rangle v\right) &= \sum_{u,v} \langle S, u\rangle\langle T, v\rangle uv \\ &= \sum_w \left(\sum_{uv=w} \langle S, u\rangle\langle T, v\rangle\right) w.\end{aligned}$$

Algebraically inclined readers can think of $K\langle\langle X\rangle\rangle$ as the completion of the monoid algebra of the free monoid X^* with respect to the ideal generated by X.

There is an obvious notion of *convergence* of a sequence $S_1, S_2, \ldots$ of formal series (and hence of a sum $\sum_{n\geq 0} T_n$) analogous to the commutative case (Section 1.1). Namely, we say that $S_1, S_2, \ldots$ converges to S if for all $w \in X^*$ the sequence $\langle S_1, w\rangle, \langle S_2, w\rangle, \ldots$ has only finitely many terms unequal to $\langle S, w\rangle$. Suppose now that $\langle S, 1\rangle = \alpha \neq 0$. Let

$$T = \frac{1}{\alpha}\sum_{n\geq 0}\left(1 - \frac{S}{\alpha}\right)^n.$$

This sum converges formally, and it is easy to check that $ST = TS = 1$. Hence $T = S^{-1}$ in $K\langle\langle X\rangle\rangle$. For instance,

$$\left(1 - \sum_{x\in X} x\right)^{-1} = \sum_{w\in X^*} w.$$

There are two subalgebras of $K\langle\langle X\rangle\rangle$ with which we will be concerned in this section.

6.5.2 Definition. (a) A (noncommutative) *polynomial* is a series $S \in K\langle\langle X\rangle\rangle$ that is a *finite* sum $\sum \langle S, w\rangle w$. The set of polynomials $S \in K\langle\langle X\rangle\rangle$ forms a subalgebra of $K\langle\langle X\rangle\rangle$ denoted $K\langle X\rangle$ (or sometimes $K_{\text{pol}}\langle\langle X\rangle\rangle$).
(b) A (noncommutative) *rational* series is an element of the smallest subalgebra $K_{\text{rat}}\langle\langle X\rangle\rangle$ of $K\langle\langle X\rangle\rangle$ containing $K\langle X\rangle$ (or equivalently, containing X) such that if $S \in K_{\text{rat}}\langle\langle X\rangle\rangle$ and S^{-1} exists, then $S^{-1} \in K_{\text{rat}}\langle\langle X\rangle\rangle$.

An example of a polynomial (when $K = \mathbb{Q}$, $X = \{x, y, z\}$) is $x^2z - xz^2 - 3x^5yxz^2 + \frac{2}{3}zyzy^2z^2$. An example of a rational series is

$$\begin{aligned}S = {} & [(1+x)^{-1} + y]^{-1}[x^2 - y^2xy(1 - xyz)^{-1} \\ & \times z(1 + xyxzx^2 + 2y^2z^3xyx)^{-1}zy^5z] + x.\end{aligned}$$

Note that notation such as $\frac{x}{1-y}$ is ambiguous; it could mean either $x(1-y)^{-1}$ or $(1-y)^{-1}x$. Note also that there does not exist a notion of "common denominator" for noncommutative series. For instance, there is no polynomial S satisfying

$$S[(1-x)^{-1}+(1-y)^{-1}] \in K\langle x, y\rangle,$$

or even polynomials S and T satisfying

$$S[(1-x)^{-1}+(1-y)^{-1}+(1-z)^{-1}]T \in K\langle x, y, z\rangle.$$

In particular, not every rational series is a quotient of two polynomials.

Let $\phi : K\langle\langle X\rangle\rangle \to K[[X]]$ be the continuous algebra homomorphism defined by $\phi(x) = x$ for all $x \in X$. Thus $\phi(S)$ is the "abelianization" of S, and the kernel of ϕ is the two-sided ideal of $K\langle\langle X\rangle\rangle$ generated by $\{xy - yx : x, y \in X\}$. Note that if $S \in K\langle X\rangle$ then $\phi(S) \in K[X]$. The converse is clearly false, e.g., if $S = \sum_{n\geq 0}(x^n y^n - y^n x^n)$ then $\phi(S) = 0 \in K[x, y]$, but $S \notin K\langle X\rangle$. Similarly if $S \in K_{\text{rat}}\langle\langle X\rangle\rangle$ then clearly $\phi(S) \in K_{\text{rat}}[[X]] = K[[X]] \cap K(X)$. Again the converse is false, but an example is not so obvious, since at this point we have no easy way to recognize when a series S is not rational (other than the condition that $\phi(S)$ is not rational). Exercise 6.65 will give us a method for showing that many series are not rational. For instance, $S = \sum_{n\geq 0} x^n y^n \notin K_{\text{rat}}\langle\langle x, y\rangle\rangle$, though $\phi(S) = 1/(1-xy) \in K_{\text{rat}}[[x, y]]$.

We next define a class of series of crucial importance in understanding rational series. We let $K^{n\times n}$ denote the monoid of all $n \times n$ matrices over K under the usual multiplication of matrices.

6.5.3 Definition. A series $S \in K\langle\langle X\rangle\rangle$ is *recognizable* if there exists a positive integer n and a homomorphism of monoids

$$\mu : X^* \to K^{n\times n},$$

as well as two matrices $\lambda \in K^{1\times n}$ and $\gamma \in K^{n\times 1}$ (so λ is a row vector and γ a column vector) such that for all $w \in X^+$ we have

$$\langle S, w\rangle = \lambda \cdot \mu(w) \cdot \gamma. \tag{6.41}$$

Note. Equation (6.41) is only required to hold for $w \in X^+$, not $w \in X^*$. In other words, the property that a series S is recognizable does not depend on the constant term $\langle S, 1\rangle$ of S.

Note. If $\lambda \neq [0, 0, \ldots, 0]$ and $\gamma \neq [0, 0, \ldots, 0]^t$ (where ^{t} denotes transpose), then it is easy to see that we can find an invertible matrix $A \in K^{n\times n}$ such that $\lambda = [1, 0, \ldots, 0]A$ and $\gamma = A^{-1}[0, 0, \ldots, 1]^t$. If we define a new homomorphism

$\mu' : X^* \to K^{n\times n}$ by $\mu'(w) = A\mu(w)A^{-1}$, then

$$\langle S, w\rangle = \mu'(w)_{1n}, \tag{6.42}$$

the $(1, n)$ entry of $\mu'(w)$. Hence we may replace (6.41) by the stronger condition (6.42).

6.5.4 Example. Suppose that $X = \{x, y\}$ and that $\mu : X^* \to K^{2\times 2}$ is defined by

$$\mu(x) = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} := a, \qquad \mu(y) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} := b.$$

The series B given by $\langle B, w\rangle = \mu(w)_{12}$ is recognizable by Definition 6.5.3. Let us see if we can obtain a formula for B. Define, for $w \in X^+$,

$$\mu(w) = \begin{bmatrix} A_w & B_w \\ C_w & D_w \end{bmatrix}, \tag{6.43}$$

and define series $A = \sum A_w w$, etc. Also set $\mu(1) = I$, the 2×2 identity matrix. Thus A_w is short for $\langle A, w\rangle$, etc., and the series B defined by (6.43) coincides with the definition of B preceding (6.43). Note that

$$\mu(wx) = \mu(w)a = \begin{bmatrix} A_w + 2B_w & B_w \\ * & * \end{bmatrix} \tag{6.44}$$

$$\mu(wy) = \mu(w)b = \begin{bmatrix} A_w & A_w + B_w \\ * & * \end{bmatrix}, \tag{6.45}$$

where $*$'s denote entries that turn out to be irrelevant. From (6.44) and (6.45) there follows

$$\begin{aligned} A_{wx} &= A_w + 2B_w, \qquad & B_{wx} &= B_w \\ A_{wy} &= A_w, & B_{wy} &= A_w + B_w. \end{aligned}$$

Hence

$$\begin{aligned} A &= 1 + \sum_w A_{wx} wx + \sum_w A_{wy} wy \\ &= 1 + \sum_w (A_w + 2B_w) wx + \sum_w A_w wy \\ &= 1 + A(x + y) + 2Bx. \end{aligned}$$

Similarly $B = Ay + B(x+y)$. Thus we obtain two linear equations in two unknowns A and B, viz.,

$$\begin{aligned} A(1-x-y) - 2Bx \qquad &= 1 \\ -Ay \qquad + B(1-x-y) &= 0. \end{aligned}$$

We now solve these equations, essentially by "noncommutative" Gaussian elimination. Since the unknowns A and B are only multiplied on the right and since the diagonal coefficients $1-x-y$ and $1-x-y$ are invertible, there will be no difficulty in carrying out the elimination. Multiply the first equation on the right by $(1-x-y)^{-1}y$ and add it to the second equation to get the following formula for B:

$$B = (1-x-y)^{-1}y[1-x-y-2x(1-x-y)^{-1}y]^{-1}. \tag{6.46}$$

Note that

$$\begin{aligned} \phi(B) &= \frac{y}{(1-x-y)\left(1-x-y-\frac{2xy}{1-x-y}\right)} \\ &= \frac{y}{(1-x-y)^2-2xy}. \end{aligned}$$

We can also compute $\phi(B)$ directly from a and b using Theorem 4.7.2. We have

$$\begin{aligned} \phi(B) &= \sum_{n\geq 0}(ax+by)^n\bigg|_{12} \\ &= (1-ax-by)^{-1}|_{12} \\ &= \begin{bmatrix} 1-x-y & -y \\ -2x & 1-x-y \end{bmatrix}^{-1}\Bigg|_{12}. \end{aligned}$$

Writing

$$c = \begin{bmatrix} 1-x-y & -y \\ -2x & 1-x-y \end{bmatrix},$$

there follows (using the notation of Theorem 4.7.2)

$$\begin{aligned} \phi(B) &= \frac{-\det[c:2,1]}{\det c} \\ &= \frac{y}{(1-x-y)^2-2xy}, \end{aligned}$$

as before.

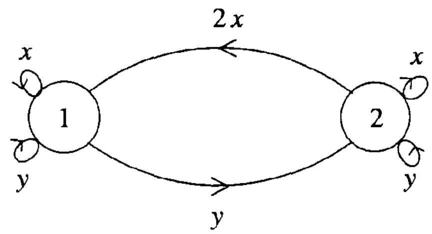


Figure 6-4. A directed graph Γ_μ.

There is an alternative graph-theoretic way to view recognizable series. Let $\mu : X^* \to K^{n\times n}$ be as in Definition 6.5.3, where we assume λ and γ are such that $\lambda \cdot \mu(w) \cdot \gamma = \mu(w)_{1n}$ as in (6.42). Define a directed graph Γ_μ with weighted edges as follows: Let $X = \{x_1, \ldots, x_r\}$, and let Γ_μ have vertex set $V = V(\Gamma_\mu) = [n]$. For each triple $(i, j, k) \in [r]^3$, draw an edge e from i to j labeled $\omega(e) = \mu(x_k)_{ij}x_k$, where $\mu(x_k)_{ij}$ denotes the (i, j) entry of the matrix $\mu(x_k)$. Thus there are r edges from vertex i to vertex j (though edges labeled 0 may be suppressed). In the case of Example 6.5.4, the digraph Γ_μ is given by Figure 6-4.

Let P be a walk in Γ_μ of length m from i to j, say $i = i_0, e_1, i_1, e_2, \ldots, i_{m-1}, e_m, i_m = j$. Define the *weight* of P by

$$\omega(P) = \omega(e_1)\omega(e_2)\cdots\omega(e_m) = \kappa(P)x_{i_1}\cdots x_{i_m},$$

a noncommutative monomial in the variables X, multiplied by some scalar $\kappa(P) \in K$. The scalar $\kappa(P)$ is just a term in the (i, j) entry of $\mu(x_{i_1})\cdots\mu(x_{i_m})$, by definition of matrix multiplication. Hence the series S defined by $\langle S, w\rangle = \mu(w)_{1n}$ is also given by

$$S = \sum_P \omega(P), \tag{6.47}$$

summed over all walks from 1 to n.

6.5.5 Example. The walks from 1 to 2 of length at most two in Figure 6-4 are given by

$$1 \xrightarrow{y} 2$$

$$1 \xrightarrow{x} 1 \xrightarrow{y} 2$$

$$1 \xrightarrow{y} 1 \xrightarrow{y} 2$$

$$1 \xrightarrow{y} 2 \xrightarrow{x} 2$$

$$1 \xrightarrow{y} 2 \xrightarrow{y} 2$$

Hence the formal series B of Example 6.5.4 begins

$$B = y + xy + yx + 2y^2 + \cdots.$$

With a little practice one can see by inspection that the series B of (6.46) is also given by (6.47). The first factor $(1-x-y)^{-1}$ in (6.46) corresponds to the initial part of the walk P before it leaves vertex 1. We can walk along the loops at vertex 1 labeled x and y in any desired order. The factor y in (6.46) corresponds to the first step from vertex 1 to vertex 2. Now we are free to walk along the loops at 2 in any order (accounting for the terms $-x-y$ in the third factor of (6.46)), then to move back to 1 (the factor $2x$ of the term $2x(1-x-y)^{-1}y$ of the third factor in (6.46)), then to walk along the loops at 1 (the factor $(1-x-y)^{-1}$ of $2x(1-x-y)^{-1}y$), then to move back to 2 (the factor y of $2x(1-x-y)^{-1}y$), and then to iterate the procedure that begins at vertex 2.

The above discussion shows that the theory of recognizable series is essentially the same as the transfer-matrix method of Section 4.7, except that we must keep track of the actual walks (i.e., the order of their edges), and not just their unordered (commutative) weights. One might say that a recognizable series is simply the generating function for (weighted) walks in a digraph. One can also view a graph such as Figure 6-4 as a kind of finite-state machine (automaton) for producing the series S. We will not say more about this point of view here, though it can be a fruitful way of looking at recognizable series.

Before stating the main theorem on rational series, we need one simple lemma for ensuring that certain series are rational.

6.5.6 Lemma. *Suppose that $B_1, \ldots, B_n$ are formal series satisfying n linear equations of the form*

$$\begin{aligned}
&B_1(1+c_{11}) + B_2c_{12} + \cdots + B_nc_{1n} = d_1\\
&B_1c_{21} + B_2(1+c_{22}) + \cdots + B_nc_{2n} = d_2\\
&\qquad\vdots\\
&B_1c_{n1} + B_2c_{n2} + \cdots + B_n(1+c_{nn}) = d_n,
\end{aligned}$$

where each c_{ij} is a rational series with zero constant term, and where each d_j is a rational series. Then $B_1, \ldots, B_n$ are rational series (and are the unique series satisfying the above system of linear equations).

Proof. Induction on n. When $n=1$ we have $B_1 = d_1(1+c_{11})^{-1}$, as desired. Now assume the result for $n-1$. Multiply the first equation on the right by $-(1+c_{11})^{-1}c_{j1}$ and add to the jth equation, for $2 \le j \le n$. We obtain a system of $n-1$ equations for $B_2, \ldots, B_n$ satisfying the hypotheses of the lemma, so $B_2, \ldots, B_n$ are rational (and unique). By symmetry (or by using the first equation to solve for B_1), we get that B_1 is also rational (and unique). $\square$

We are now ready to state the main theorem on rational series.

6.5.7 Theorem (Fundamental theorem of rational formal series). *A formal series $S \in K\langle\langle X\rangle\rangle$ is rational if and only if it is recognizable.*

Proof (sketch). Assume S is recognizable. The proof that S is rational parallels the computation of Example 6.5.4. Let $\mu : X^+ \to K^{n\times n}$ be a homomorphism of monoids satisfying $\langle S, w\rangle = \mu(w)_{1n}$ for all $w \in X^+$. Set $\mu(1) = I$, the $n \times n$ identity matrix. Define series

$$A_{ij} = \sum_{w\in X^+} \mu(w)_{ij} w,$$

for $(i, j) \in [n] \times [n]$. If $x_k \in X$, then let $a^k = \mu(x_k)$ and $a^k_{ij} = \mu(x_k)_{ij}$, the (i, j) entry of the matrix a^k. Then $\mu(wx_k) = \mu(w)a^k$, so

$$\mu(wx_k)_{ij} = \sum_t \mu(w)_{it} a^k_{tj}.$$

Multiplying by wx_k and summing on w and k yields when $i = 1$ the equations

$$A_{1j} = \delta_{1j} + \sum_t \sum_k a^k_{tj} A_{1t} x_k,$$

or equivalently

$$\sum_t A_{1t}\left(\delta_{tj} - \sum_k a^k_{tj} x_k\right) = \delta_{1j}, \qquad 1 \le j \le n. \tag{6.48}$$

This is a system of n linear equations in the n unknowns $A_{11}, A_{12}, \ldots, A_{1n}$, which has the form given by Lemma 6.5.6. Hence A_{1n} (as well as $A_{11}, A_{12}, \ldots, A_{1,n-1}$) is rational, as desired.

To show conversely that rational series are recognizable, we must show the following four facts.

(i) If $x_i \in X$, then the series x_i is recognizable.
(ii) If S and T are recognizable and $\alpha, \beta \in K$, then $\alpha S + \beta T$ is recognizable.
(iii) If S and T are recognizable, then ST is recognizable.
(iv) If S is recognizable and $\langle S, 1\rangle \neq 0$, then S^{-1} is recognizable.

Fact (i) is trivial, while the remaining three facts are all proved by constructing appropriate homomorphisms $X^* \to K^{n\times n}$. The details are tedious and will not be given here. □

6.6 Algebraic Formal Series

In this section we consider an important generalization of the class of rational series. As before, we have a fixed (finite) alphabet X.

6.6.1 Definition. Let $Z = \{z_1, \ldots, z_n\}$ be an alphabet disjoint from X. A *proper algebraic system* is a set of equations $z_i = p_i$, $1 \le i \le n$, where:

(a) $p_i \in K\langle X, Z\rangle$ (i.e., p_i is a polynomial in the alphabet $X \cup Z$);
(b) $\langle p_i, 1\rangle = 0$ and $\langle p_i, z_j\rangle = 0$ (i.e., p_i has no constant term and no terms $c_j z_j$, $0 \neq c_j \in K$).

We sometimes also call the n-tuple $(p_1, \ldots, p_n)$ a *proper algebraic system*.

A *solution* (sometimes called a *strong solution*) to a proper algebraic system $(p_1, \ldots, p_n)$ is an n-tuple $(R_1, \ldots, R_n) \in K\langle\langle X\rangle\rangle^n$ of formal series in X with zero constant term satisfying

$$R_i = p_i(X, Z)_{z_i = R_i}. \tag{6.49}$$

(The series $p_i(X, Z)_{z_i=R_i}$ will be formally well defined, since $\langle R_i, 1\rangle = 0$ by assumption.) Each R_i is called a *component* of the system $(p_1, \ldots, p_n)$.

6.6.2 Example. If $X = \{x, y\}$ and $Z = \{z\}$, then

$$z = xy + xzy$$

is a proper algebraic system with solution

$$R = \sum_{n\geq 1} x^n y^n. \tag{6.50}$$

Compare the system $z = xy + zxy$, with solution $R = \sum_{n\geq 1}(xy)^n = (1-xy)^{-1} - 1$, a rational series. It can be shown that the series (6.50) is not rational (see Exercise 6.65).

6.6.3 Proposition. *Every proper algebraic system $(p_1, \ldots, p_n)$ has a unique solution $R = (R_1, R_2, \ldots, R_n) \in K\langle\langle X\rangle\rangle^n$.*

Idea of Proof. The method of proof is "successive approximation." Define the first approximation $S^{(1)} = (S_1^{(1)}, \ldots, S_n^{(1)})$ to a solution by $S_i^{(1)} = p_i(X, Z) \in K\langle X, Z\rangle$. Now assuming that the k-th approximation $S^{(k)} = (S_1^{(k)}, \ldots, S_n^{(k)}) \in K\langle X, Z\rangle^n$ has been defined, let

$$S_i^{(k+1)} = p_i\big(X, S^{(k)}\big)$$

(i.e., substitute $S_j^{(k)}$ for Z_j in $p_i(X, Z)$). It is straightforward to verify from the definition of proper algebraic system that $\lim_{k\to\infty} S^{(k)}$ converges formally to a solution $R \in K\langle\langle X\rangle\rangle^n$, and that this solution must be unique.

As a simple example with just one equation, consider the proper algebraic system

$$z = x + xzy + yz.$$

The first approximation is

$$S^{(1)} = x + xzy + yz.$$

The second approximation is

$$\begin{aligned} S^{(2)} &= x + x(x + xzy + yz)y + y(x + xzy + yz) \\ &= x + x^2y + x^2zy^2 + xyzy + yx + yxzy + y^2z. \end{aligned}$$

The third approximation is

$$\begin{aligned} S^{(3)} &= x + xS^{(2)}y + yS^{(2)} \\ &= x + x^2y + x^3y^2 + xyxy + yx + yx^2y + y^2x + \text{ terms involving } z. \end{aligned}$$

This last approximation agrees with the solution R in all terms of degree at most three (and with some nonzero terms of higher degree).

NOTE (for logicians). Let $(p_1, \dots, p_m)$ and $(p'_1, \dots, p'_n)$ be proper algebraic systems with solutions $(R_1, \dots, R_m)$ and $(R'_1, \dots, R'_n)$, respectively. It is undecidable whether $R_1 = R'_1$, or in particular whether $R_1 = 0$. (One difficulty is that substituting $z_1 = 0$ in the system $z_1 = p_1, \dots, z_m = p_m$ may yield a system which isn't proper.) On the other hand, it *is* decidable whether the abelianizations $\phi(R_1)$ and $\phi(R'_1)$ are equal.

6.6.4 Definition. (a) A series $S \in K\langle\langle X\rangle\rangle$ is *algebraic* if $S - \langle S, 1\rangle$ is a component of a proper algebraic system. The set of all algebraic series $S \in K\langle\langle X\rangle\rangle$ is denoted $K_{\text{alg}}\langle\langle X\rangle\rangle$.
(b) The *support* of a series $S = \sum \langle S, w\rangle w \in K\langle\langle X\rangle\rangle$ is defined by

$$\operatorname{supp}(S) = \{w \in X^* : \langle S, w\rangle \neq 0\}.$$

A *language* is a subset of X^*. A language L is said to be *rational* (respectively, *algebraic*) if it is the support of a rational (respectively, algebraic) series. A rational language is also called *regular*, and an algebraic language is also called *context-free.*

Our previous example (Example 6.6.2) of a proper algebraic system yields the algebraic series $\sum x^n y^n$ and $\sum (xy)^n$. Let us consider some examples of greater interest.

6.6.5 Example. Rational series are algebraic. The system (6.48) of linear equations is equivalent to a proper algebraic system in the unknowns $A_{11} - 1, A_{12}, \dots, A_{1n}$.

6.6.6 Example. The *Dyck language* D is the subset of $\{x, y\}^*$ such that if x is replaced by a left parenthesis and y by a right parenthesis, then we obtain a sequence of properly nested parentheses. Equivalently, a word $w_1 w_2 \cdots w_m$ is in D (where $w_i = x$ or y) if for all $1 \leq j \leq m$ the number of x's among $w_1 w_2 \cdots w_j$ is at least as great as the number of y's among $w_1 w_2 \cdots w_j$, and the total number of x's is equal to the total number of y's (so m is even). An element w of D is

called a *Dyck word.* The Dyck words of length six or less are given by

$$1 \quad xy \quad x^2y^2 \quad xyxy \quad x^3y^3 \quad x^2yxy^2 \quad x^2y^2xy \quad xyx^2y^2 \quad xyxyxy.$$

By Corollary 6.2.3(ii) or (iii), it follows that the number of words of length $2n$ in D is given by the Catalan number C_n. Now note the following key recursive property of a Dyck word w: If w is nonempty, then it begins with an x, followed by a Dyck word, then by a y (the right parenthesis matching the initial x), and finally by another Dyck word. Thus D is a solution to the system

$$z = 1 + xzyz,$$

and so $D^+ = D - 1$ is a solution to the proper algebraic system

$$\begin{aligned} z' &= x(z'+1)y(z'+1) \\ &= xy + xyz' + xz'y + xz'yz'. \end{aligned} \tag{6.51}$$

It follows from (6.51) that the Dyck language D is algebraic. It is perhaps the most important algebraic language for enumerative combinatorics. Many enumerative problems can be expressed in terms of the Dyck language, e.g., many of the parts of Exercise 6.19.

6.6.7 Example. Let $X = \{x_0, x_1, \ldots, x_m\}$. Define a weight $\omega : X \to \mathbb{Z}$ by $\omega(x_i) = i - 1$. Define a language $L \subset X^*$ by

$$\begin{aligned} L = \{x_{i_1}x_{i_2}\cdots x_{i_k} : \omega(x_{i_1}) + \cdots + \omega(x_{i_j}) \geq 0 \text{ if } j < k, \\ \text{and } \omega(x_{i_1}) + \cdots + \omega(x_{i_k}) = -1\}. \end{aligned}$$

Thus L consists of all words that encode plane trees of maximum degree m, as discussed in Section 5.3 (see Lemma 5.3.9). The language L is called the *Łukasiewicz language*, and its elements are *Łukasiewicz words.* Note that a Łukasiewicz word in the letters x_0 and x_2 is just a Dyck word with x replaced by x_2 and y by x_0, with an x_0 appended at the end. It is easily verified by inspection (essentially the definition of a plane tree) that L^+ is a solution (or component) of the proper algebraic system

$$z = x_0 + x_1 z + x_2 z^2 + \cdots + x_m z^m.$$

Hence the Łukasiewicz language is algebraic.

The next example will be of use to us later (Theorem 6.7.1) when we consider noncommutative diagonals. It is a good example of the nonobvious way in which auxiliary series may need to be introduced in order to obtain a proper algebraic system for which some given series S is a component.

6.6.8 Example. Let $X = \{x_1, \ldots, x_k, y_1, \ldots, y_k\}$, and let $\Delta \subset X^*$ be the set of those elements of X^* that reduce to the identity under the relations

$$x_i y_i = y_i x_i = 1, \qquad 1 \leq i \leq k.$$

In other words, if in the word $w \in X^*$ we replace y_i by x_i^{-1}, then we obtain the identity element of the free group generated by $x_1, \ldots, x_k$. Thus for example when $k = 3$ we have

$$x_1^2\, y_2^2\, x_2\, x_3\, y_3\, x_2\, y_1\, x_2\, y_2\, y_1 \in \Delta.$$

We claim that Δ is algebraic. If $t \in X$ let

$$G_t = \{w \in \Delta : w = tv,\ w \neq uu' \text{ for } u, u' \in \Delta^+\}.$$

Write

$$\bar{t} = \begin{cases} x_i & \text{if } t = y_i \\ y_i & \text{if } t = x_i. \end{cases}$$

Any word $w \in G_t$ must end in $\bar{t}$ [why?], so we can define a series (or language) B_t by $G_t = tB_t\bar{t}$. It is then not difficult to verify that

$$\begin{aligned} \Delta &= 1 + \Delta \sum_{t \in X} G_t \\ G_t &= tB_t\bar{t} \\ B_t &= 1 + B_t \sum_{\substack{q \in X \\ q \neq \bar{t}}} G_q. \end{aligned}$$

From these formulas we see that

$$\begin{aligned} \Delta^+ &= (\Delta^+ + 1) \sum_{t \in X} t\big(B_t^+ + 1\big)\bar{t} \\ B_t^+ &= \big(B_t^+ + 1\big) \sum_{\substack{q \in X \\ q \neq \bar{t}}} q\big(B_q^+ + 1\big)\bar{q}, \qquad t \in X. \end{aligned}$$

Hence $(\Delta^+, (B_t^+)_{t \in X})$ is a solution to a proper algebraic system, so Δ is algebraic.

We now want to relate algebraic formal series to *commutative* algebraic generating functions as discussed in Sections 6.1–6.3. We need a standard result in the theory of extension fields, which we simply state without proof.

6.6.9 Lemma. *Suppose that $\alpha_1, \ldots, \alpha_n$ belong to an extension field of a field K. Suppose also that there exist polynomials $f_1, \ldots, f_n \in K[X]$, where $X = (x_1, \ldots, x_n)$, satisfying:*

(i) $f_i(\alpha_1, \ldots, \alpha_n) = 0, \quad 1 \leq i \leq n,$
(ii) $\det(\partial f_i / \partial \alpha_j) \neq 0.$

Then each α_i is algebraic (in fact, separably algebraic, though separability is irrelevant here) over K.

Let us give a couple of examples showing the significance of the conditions (i) and (ii) above. If $f_1 = x_1 - x_2$ and $f_2 = (x_1 - x_2)^2$, then condition (i) is satisfied for

any $\alpha_1 = \alpha_2$, but $\det(\partial f_i/\partial \alpha_j) = 0$. In fact, the Jacobian determinant $\det(\partial f_i/\partial x_j)$ is identically zero, since the polynomials f_1 and f_2 are algebraically dependent. Now let $f_1 = x_1$, $f_2 = x_1x_2$, so f_1 and f_2 are algebraically independent and $\det(\partial f_i/\partial x_j) = x_1 \neq 0$. The solutions (α_1, α_2) to $f_1 = f_2 = 0$ are $(0, \alpha_2)$ for *any* α_2, so α_2 need not be algebraic over K. This does not contradict Lemma 6.6.9, since $\det(\partial f_i/\partial \alpha_j) = 0$.

6.6.10 Theorem. *Let $S \in K_{\text{alg}}\langle\langle X\rangle\rangle$, where X is a finite alphabet. Then $\phi(S)$ is algebraic over the field $K(X)$ of rational functions in (the commuting variables) X.*

Proof. We can assume $\langle S, 1\rangle = 0$. Let $(p_1, \ldots, p_n)$ be a proper algebraic system with solution $(S_1 = S, S_2, \ldots, S_n)$. Let $\eta_i = \phi(S_i)$ and $\eta = (\eta_1, \ldots, \eta_n)$. Let $Z = (z_1, \ldots, z_n)$ and

$$f_i(Z) = z_i - p_i(X, Z).$$

Hence, regarding the p_i as *commutative* polynomials, the η_i satisfy

$$f_i(\eta) = \eta_i - p_i(X, \eta) = 0.$$

The Jacobian matrix at $Z = \eta$ is given by

$$J = \left(\frac{\partial f_i}{\partial \eta_j}\right) = I - \left(\frac{\partial p_i}{\partial \eta_j}\right).$$

By definition of proper algebraic system (and because $\langle S_i, 1\rangle = 0$), all entries $\partial p_i/\partial \eta_j$ belong to the maximal ideal $XK[[X]] = x_1K[[X]] + \cdots + x_kK[[X]]$ of the ring $K[[X]]$. Hence $\det J = 1 + m$, where $m \in XK[[X]]$. Thus $\det J \neq 0$, so by Lemma 6.6.9 the η_i's are algebraic over $K(X)$. □

6.6.11 Example. Let D be the Dyck language of Example 6.6.6. We saw in that example that D is algebraic. If C_n is the number of Dyck words of length $2n$ (which we know is a Catalan number) then $\phi(D) = \sum_{n\geq 0} C_n x^n y^n$. Hence by Theorem 6.6.10 (substituting x for xy) we get that $\sum_{n\geq 0} C_n x^n$ is algebraic directly from the combinatorial structure of the Dyck language.

Our final topic in this section will be the Hadamard product. If $S = \sum \langle S, w\rangle w$ and $T = \sum \langle T, w\rangle w$ are two formal series (over the same alphabet X), then define, just as in the commutative case, the *Hadamard product*

$$S * T = \sum \langle S, w\rangle \langle T, w\rangle w.$$

(The notation $S \odot T$ is also used.) We will need the following result in the next section. It is the noncommutative analogue of Proposition 6.1.11. (For a related result, see Exercise 6.66).

6.6.12 Proposition. *If $S \in K_{\text{rat}}\langle\langle X\rangle\rangle$ and $T \in K_{\text{alg}}\langle\langle X\rangle\rangle$, then $S * T \in K_{\text{alg}}\langle\langle X\rangle\rangle$.*

Proof (sketch). Let $z_i = p_i$, $1 \le i \le n$, be a proper algebraic system with solution $z_1 = T$. Since S is rational, it is recognizable by Theorem 6.5.7. Hence for some $m \ge 1$ there is a monoid homomorphism $\mu : X^* \to K^{m\times m}$ such that

$$\langle S, w\rangle = \mu(w)_{1m} \qquad \text{for all } w \in X^+.$$

Introduce nm^2 new variables z_i^{jk}, $(i, j, k) \in [n] \times [m] \times [m]$. Let $Z' = \{z_i^{jk} : (i, j, k) \in [n] \times [m] \times [m]\}$, and let M_i be the $m \times m$ matrix with (j, k) entry given by $(M_i)_{jk} = z_i^{jk}$. For each $(i, j, k) \in [n] \times [m] \times [m]$, construct a polynomial $p_i^{jk} \in K\langle X \cup Z'\rangle$ as follows: Replace in p_i each letter z_t by the matrix M_t and each letter $x \in X$ by the matrix $\mu(x)x$ (scalar multiplication of $\mu(x)$ by x). Performing all matrix additions and multiplications involved in p_i transforms p_i into an $m \times m$ matrix. Let p_i^{jk} be the (j, k) entry of this matrix.

Now define a system $\mathcal{S}$ by

$$z_i^{jk} = p_i^{jk}, \qquad (i, j, k) \in [n] \times [m] \times [m].$$

Clearly $\mathcal{S}$ is proper. Moreover, it's not hard to check that if the original system $z_i = p_i$ has the solution $(z_1, \ldots, z_n) = (t_1 = T, t_2, \ldots, t_n)$, then for all $w \in X^*$ the solution $(z_i^{jk}) = (s_i^{jk})$ to $\mathcal{S}$ satisfies

$$\left\langle s_i^{jk}, w\right\rangle = \mu(w)_{jk}\langle t_i, w\rangle.$$

In particular,

$$\begin{aligned}\left\langle s_1^{1m}, w\right\rangle &= \mu(w)_{1m}\langle t_1, w\rangle\\ &= \langle S, w\rangle\langle T, w\rangle.\end{aligned}$$

Hence $S * T \in K_{\text{alg}}\langle\langle X\rangle\rangle$. □

6.6.13 Example. Let $X = \{x_1, x_2\}$, and let

$$\mu(x_1) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} := a, \qquad \mu(x_2) = \begin{bmatrix} -1 & 0 \\ 2 & 1 \end{bmatrix} := b.$$

It's easy to check that $(a^r b^r)_{12} = r$ for all $r \ge 1$, so the series S of the previous proposition satisfies $\langle S, x_1^r x_2^r\rangle = r$. Define T to be the solution to the proper algebraic system

$$z = x_1 z x_2 - x_1 x_2,$$

so $T = -\sum_{r\ge1} x_1^r x_2^r$ and $S * T = -\sum_{r\ge1} r x_1^r x_2^r$. The system $\mathcal{S}$ of the previous proof is defined by

$$\begin{bmatrix} z^{11} & z^{12} \\ z^{21} & z^{22} \end{bmatrix} = \begin{bmatrix} x_1 & x_1 \\ 0 & x_1 \end{bmatrix}\begin{bmatrix} z^{11} & z^{12} \\ z^{21} & z^{22} \end{bmatrix}\begin{bmatrix} -x_2 & 0 \\ 2x_2 & x_2 \end{bmatrix} - \begin{bmatrix} x_1 & x_1 \\ 0 & x_1 \end{bmatrix}\begin{bmatrix} -x_2 & 0 \\ 2x_2 & x_2 \end{bmatrix}.$$

In other words,

$$z^{11} = -x_1 z^{11} x_2 - x_1 z^{21} x_2 + 2x_1 z^{12} x_2 + 2x_1 z^{22} x_2 - x_1 x_2,$$

and similarly for z^{21}, z^{12}, z^{22}. The assertion of the previous proof is that the z^{12}-component is given by $s^{12} = -\sum_{r\ge 1} r x_1^r x_2^r$.

6.7 Noncommutative Diagonals

In this section we will use the theory of noncommutative formal series to show that certain (ordinary) Laurent series $\eta \in K((t))$ are algebraic. We will be dealing with series of the form

$$S(x_1, \dots, x_k, y_1, \dots, y_k) \in K(t)_{\mathrm{rat}}\langle\langle X, Y\rangle\rangle, \tag{6.52}$$

where $X = (x_1, \dots, x_k)$, $Y = (y_1, \dots, y_k)$. In other words, S is a rational series in X and Y with coefficients in the field $K(t)$ of rational functions in one variable t (commuting with X and Y). Since $K(t) \subset K((t))$ we can regard the coefficients of S as Laurent series in t. We will assume that S is such that the series $S(X, X^{-1}) := S(x_1, \dots, x_k, x_1^{-1}, \dots, x_k^{-1})$ is well-defined, i.e., the coefficient of a Laurent monomial $x_{i_1}^{a_1} \cdots x_{i_j}^{a_j}$ $(a_1, \dots, a_j \in \mathbb{Z})$ is a formal Laurent series $\eta \in K((t))$ in the variable t.

For instance (writing $x = x_1$ and $y = y_1$), if

$$S(x, y) = \sum_{n\ge 0} t^n (xy)^n = (1 - txy)^{-1},$$

then

$$S(x, x^{-1}) = \sum_{n\ge 0} t^n.$$

This is certainly a well-defined series; the coefficient of x^i for $i \ne 0$ is 0, while the coefficient of x^0 is the Laurent series (in fact, power series) $\sum_{n\ge 0} t^n$. For a more complicated example, let

$$\begin{aligned} S(X, Y) &= \sum_{n\ge 0} t^n (x_1 + \cdots + x_k + y_1 + \cdots + y_k)^n \\ &= [1 - t(x_1 + \cdots + x_k + y_1 + \cdots + y_k)]^{-1}. \end{aligned}$$

Fix a noncommutative Laurent monomial u in X, e.g., $u = x_1^{-2} x_2 x_3^{-1} x_1 x_3^2$. Then

$$[u]S(X, X^{-1}) = \sum_{n\ge 0} t^n [u]\big(x_1 + \cdots + x_k + x_1^{-1} + \cdots + x_k^{-1}\big)^n.$$

Since $[u](x_1 + \cdots + x_k + x_1^{-1} + \cdots + x_k^{-1})^n$ is simply an integer, we see that $[u]S(X, X^{-1})$ is a well-defined Laurent series for every u, so $S(X, X^{-1})$ is well defined. For an example of series $S(X, Y)$ for which $S(X, X^{-1})$ *isn't* well-defined, take for instance

$$S(x, y) = \sum_{n\ge 0} (xy)^n = (1 - xy)^{-1}.$$

Note that if $f(x, y)$ is a *commutative* power series, then

$$[x^0]f(x, tx^{-1}) = (\operatorname{diag} f)(t).$$

Hence, returning to the noncommutative case, we see that $[u]S(X, X^{-1})$ (where u is a Laurent monomial in $x_1, \dots, x_k$) is a kind of "generalized noncommutative diagonal."

The main result of this section is given by the following theorem.

6.7.1 Theorem. *Let $S = S(X, Y)$ be given by (6.52). Suppose that $S(X, X^{-1})$ is well-defined formally, so for every Laurent monomial u in the variables X we have that $[u]S(X, X^{-1}) \in K((t))$. Then $[u]S(X, X^{-1}) \in K_{\text{alg}}((t))$, i.e., the Laurent series $[u]S(X, X^{-1})$ is algebraic over $K(t)$.*

Proof. Since

$$[u]S(X, X^{-1}) = [1]T(X, X^{-1})$$

where $T = u^{-1}S$, and since T is clearly rational when S is, we may assume $u = 1$. Let $\Delta = \Delta(X, Y)$ be as in Example 6.6.8. Thus the coefficient we want is the sum (which we are assuming exists formally) of all the coefficients of the Hadamard product $S * \Delta$, and hence also of the abelianization $\phi(S * \Delta)$. Since $S \in K(t)_{\text{rat}}\langle\langle X, Y\rangle\rangle$ and $\Delta \in K_{\text{alg}}\langle\langle X, Y\rangle\rangle \subset K(t)_{\text{alg}}\langle\langle X, Y\rangle\rangle$, it follows from Proposition 6.6.12 that $S * \Delta \in K(t)_{\text{alg}}\langle\langle X, Y\rangle\rangle$. Thus by Theorem 6.6.10, $\phi(S * \Delta)$ is algebraic over $K(t)((X, Y))$. By Proposition 6.1.12 it follows that $\phi(S * \Delta)_{x_i = y_i = 1}$ is algebraic over $K(t)$, as was to be proved. □

6.7.2 Corollary. *Let P be a noncommutative Laurent polynomial over K in the variables $X = (x_1, \dots, x_k)$. (Equivalently, $P \in K[F_k]$, the group algebra of the free group F_k generated by X.) Let u be a noncommutative Laurent monomial in X (i.e., $u \in F_k$). Then the power series*

$$y = \sum_{n \geq 0} ([u]P^n)t^n \in K[[t]]$$

is algebraic (over $K(t)$).

Proof. We have

$$y = [u](1 - Pt)^{-1}.$$

The proof follows from Theorem 6.7.1. □

Note that the commutative analogue of Corollary 6.7.2 fails. For instance, if $P = x + x^{-1} + y + y^{-1} \in \mathbb{C}[x, x^{-1}, y, y^{-1}]$, then

$$\sum_{n \geq 0} ([1]P^n)t^n = \sum_{m \geq 0} \binom{2m}{m}^2 t^{2m},$$

which according to Exercise 6.3 is not algebraic. Thus in the context of diagonals, we see that noncommutative series behave better than the commutative ones.

6.7.3 Example. An interesting special case of Corollary 6.7.2 is when $P = x_1 + x_2 + \cdots + x_k + x_1^{-1} + x_2^{-1} + \cdots + x_k^{-1}$ and $u = 1$. Then

$$\begin{aligned} y &= \sum_{n\geq 0} [1](x_1 + \cdots + x_k + x_1^{-1} + \cdots + x_k^{-1})^n t^n \\ &= 1 + 2kt^2 + (8k^2 - 2k)t^4 + (40k^3 - 24k^2 + 4k)t^6 + \cdots. \end{aligned}$$

It would be extremely tedious to compute y explicitly using the method inherent in the proof of Theorem 6.7.1, but a direct combinatorial argument can be used to show that

$$y = \frac{2k-1}{k-1+k\sqrt{1-4(2k-1)t^2}}.$$

See Exercise 6.74.

As an application of Theorem 6.7.1, we give another proof that the diagonal of a (commutative) rational series in two variables is algebraic (Theorem 6.3.3).

Second Proof of Theorem 3.3. By Exercise 4.1(b), we can write $F(s,t) = P(s,t)/Q(s,t)$, where $P, Q \in K[s,t]$ and $Q(0,0) \neq 0$. Define a noncommutative series $\tilde{F}(\tilde{s}, \tilde{t})$ by

$$\tilde{F}(\tilde{s}, \tilde{t}) = P(\tilde{s}, x\tilde{t})Q(\tilde{s}, x\tilde{t})^{-1} \in K(x)_{\mathrm{rat}}\langle\langle \tilde{s}, \tilde{t}\rangle\rangle.$$

The coefficient of $\tilde{s}^0$ in $\tilde{F}(\tilde{s}, \tilde{s}^{-1})$ is the sum of the coefficients in $\tilde{F}(\tilde{s}, \tilde{t})$ of monomials $w \in \{\tilde{s}, \tilde{t}\}^*$ having equal total degree in $\tilde{s}$ and $\tilde{t}$. The contribution for $\deg \tilde{s} = \deg \tilde{t} = n$ is just $f(n,n)x^n$, where $F(s,t) = \sum f(i,j)s^i t^j$. Therefore the coefficient of $\tilde{s}^0$ in $\tilde{F}(\tilde{s}, \tilde{s}^{-1})$ is just diag F. Moreover, any coefficient $[\tilde{s}^i]\tilde{F}(\tilde{s}, \tilde{s}^{-1})$ is well-defined, so the proof follows from Theorem 6.7.1. □

Notes

The theory of algebraic functions is a vast subject, but only a small part of it has been found to have direct relevance to enumerative combinatorics. It would be interesting to see whether some of the deeper aspects of algebraic functions, such as the Riemann–Roch theorem or the theory of abelian integrals, can be applied to enumerative combinatorics. The first result in our presentation that is not a simple consequence of an introductory algebra course is Puiseux's theorem (Theorem 6.1.5). First proved by V. Puiseux **[54]** in 1850, some expositions of the proof appear in [**12**, Ch. 4.6][**14**, pp. 373–396][**21**, Ch. III.6][**47**, Ch. V, Thm. 3.1][**74**, Ch. IV, Thm. 3.1]. A "modern" proof was given by P. M. Cohn **[17][18]**. For computational aspects see **[37]**.

Several interesting results concerning algebraic functions and related to enumeration appear in the paper **[40]** of R. Jungen. In particular, there is a proof of Proposition 6.1.11 and a determination of the asymptotic behavior of the coefficients of an algebraic power series. This latter result is a useful tool for showing that certain series are not algebraic (see Exercise 6.3). For a wealth of further information on the fascinating subject of discriminants, see **[28]**. Equation (6.7) is discussed further in Exercise 6.8(a).

* The enumeration of trees, parenthesizations (or bracketings), ballot sequences, lattice paths, and polygon dissections, and the close connections among them, as summarized by Proposition 6.2.1, goes back to Segner and Euler in 1760.* Segner **[67]** obtained a recurrence for the number of triangulations of a polygon (the problem discussed in our Corollary 6.2.3(vi)), and Euler **[23]** essentially solved this recurrence, though without details of a proof. (Euler published his result as an unsigned summary of the work of Segner, but it is evident that Euler is indeed the author.) The more general problem of computing the number of dissections of an n-gon by a fixed number m of its diagonals was posed by Pfaff to N. von Fuss, who generalized Segner's recurrence **[26]**. (For more information on this problem of Pfaff and Fuss, see Exercise 6.33(c).) In the period 1838–1839 four authors considered the Euler–Segner problem. The first was G. Lamé **[44]**. Lamé's proof of the Euler–Segner result was further developed and discussed by the Belgian mathematician Eugène Charles Catalan (1814–1894) **[10]**, who wrote several other papers on this topic. The other two authors were O. Rodrigues **[56][57]** and J. Binet **[7][8]**. The term "Catalan number" arose from a citation by Netto **[1.14**, §122, §124], who attributed the problems of binary parenthesization and polygon triangulation to Catalan. A good historical discussion is given in **[9]**. An extensive bibliography (up to 1976) of Catalan numbers appears in **[30]**. The thesis **[42]** contains 31 combinatorial structures enumerated by Catalan numbers and 158 bijections among them. A very readable popular exposition of Catalan numbers appears in **[27]**, while a recent survey aimed at a more mathematical audience is given by **[38]**. An earlier survey is **[2]**. For further information on Catalan numbers, see Exercises 6.19–6.36. For some interesting recent work on triangulations of polygons, see **[1][20][46][69]**. A little-known historical aspect of Catalan numbers is their independent discovery in China, beginning with Ming An-tu (1692?–1763?). He was a Mongolian mathematician who obtained several recurrences for Catalan numbers in the 1730s, though his work was not published until 1839. Ming and his successors, however, did not obtain combinatorial interpretations of Catalan numbers. For further information, see **[50]**. Recently there have been efforts to bring Catalan numbers into undergraduate and even secondary-school education; see for instance the papers **[15][39][41][73]**.

The connection between bracketings and plane trees (Proposition 6.2.1(i) and (ii)) was known to Cayley **[11]**. The bijection with polygon dissections (Proposition 6.2.1(vi)) appears in **[5.22]** (with a sequel by Erdélyi and Etherington in **[5.20]**). Ballot sequences were first considered by J. Bertrand **[6]** in 1887. He sketched a proof by induction of a ballot theorem that includes Corollary 6.2.3(ii). A famous proof based on the "reflection principle" was given soon after by D. André. See Exercise 6.20 for further details. Additional information on ballot problems appears in Chapter 7 (see Proposition 7.10.3(c) and Corollary 7.21.6).

The formula (6.18) for the number of plane trees with $(k-1)n+1$ endpoints and every internal vertex of degree k is a special case of Theorem 5.3.10. See

*See the discussion below of Schröder's second problem for a remarkable earlier reference to a special bracketing problem.

the Notes to Chapter 5 for references. The power series solution to $y^5 + y = x$ (equivalent to (6.19) in the case $k = 5$) was obtained by Eisenstein **[22]** in his work on quintic equations. See **[51]** for an interesting historical discussion. The generating functions of Example 6.2.7, as well as some related ones, are considered by Pólya in **[53]**. Pólya mentions that Hurwitz posed the problem of showing that $\sum_n \binom{\beta n}{n} x^n$ is algebraic for $\beta \in \mathbb{Q}$. The "four combinatorial problems" (*vier combinatorische Probleme*) of Schröder appear in **[5.60]**. Much additional work related to Schröder's problems has been carried out. The first problem (equivalent to triangulations of a polygon) has already been discussed. The second problem (equivalent to arbitrary dissections of a polygon) is discussed further in Exercise 6.39. The term "Schröder number" seems to have been first used by Rogers **[58]**. The third and fourth problems are discussed in the Notes to Chapter 5. In 1994 D. Hough, while a graduate student at George Washington University, made a remarkable historical discovery related to the second problem of Schröder. He observed that the mysterious number 103,049 of Exercise 1.45 is just the tenth Schröder number $s_2(10)$. In other words, Hipparchus was aware of the Schröder numbers in the second century B.C. (Should they now be called Hipparchus numbers?) This discovery solves what is perhaps the oldest open problem related to combinatorics and shows that the ancient Greeks (or at least Hipparchus) were much more sophisticated in combinatorics than previously realized. The number 310,952 of Exercise 1.45 remains an enigma, though a possible interpretation of the nearby number 310,954 has been given by L. Habsieger, M. Kazarian, and S. Lando **[32]**. For further information related to Hough's discovery, see **[71]**.

The main result of Section 6.3, that the diagonal of a rational function of two variables is algebraic (Theorem 6.3.3), is due to H. Furstenburg **[25]**. His proof was based on contour integration as in the proof sketched preceding Example 6.3.5. (For further aspects of Furstenburg's paper, see Exercise 6.11. For a rigorous discussion of the contour integration technique for computing diagonals of power series in two variables, see **[36]**.) Our first proof of Theorem 6.3.3 follows Gessel **[29**, Thm. 6.1]. A proof based on noncommutative formal series, similar to the proof we give at the end of Section 6.7, was given by M. Fliess **[24**, Prop. 5]. For more information on the Delannoy numbers of Example 6.3.8, see **[2.3**, Exercise I.21].

D-finite power series and P-recursive functions were first systematically investigated in **[70]**, though much was known about them before **[70]** appeared. D-finite series are also called *holonomic* and are involved in much recent work dealing with algorithms for discovering and verifying combinatorial identities **[52][75]**. The basic connection between D-finite series and P-recursive functions (Proposition 6.4.1) is alluded to in **[40**, p. 299]. The earliest explicit statement of the fact that algebraic functions are D-finite (Theorem 6.4.6) of which we are aware is due to Comtet **[19]**. The extension of the theory of D-finite series to several variables is discussed in **[33][34][48][49]** and references given there. Additional references on D-finiteness may be found in Exercises 6.53–6.62.

The hierarchy rational $\Rightarrow$ algebraic $\Rightarrow$ D-finite can be further extended, though these extensions have not yet proved as useful in combinatorics as the original

three classes. Two classes that may warrant further investigation are differentially finite algebraic series and constructible differentially finite algebraic series. For further information see [3][4][59][60].

The theory of rational formal noncommutative power series originated in the pioneering work of M. P. Schützenberger [64][65][66]. In particular, the Fundamental Theorem of Rational Formal Series (Theorem 6.5.7) appears in [66]. A similar theory of algebraic formal power series is due to Chomsky and Schützenberger [13]. For a comprehensive account of the theory of rational series, see [5]. Some good references for noncommutative series in general and their connections with languages and automata are [16][43][55][61][62][63][68]. These last seven references contain many additional historical remarks and references. For further information related to decidability aspects of noncommutative series, see e.g. [**43**, §8 and §16] [**61**, Ch. VIII][**63**, Chs. II.12 and IV.5]. An interesting survey of the connection between algebraic series and combinatorics appears in [**72**].

The Dyck language of Example 6.6.6 plays a fundamental role in a theorem of Chomsky and Schützenberger on the structure of arbitrary algebraic series. See [**63**, Thm. IV.4.5] for an exposition. Example 6.6.8 is due to Chomsky and Schützenberger [**13**] and is also discussed in [**35**, Prop. 3.2]. For a proof of Lemma 6.6.9, see [**45**, Chap. X, Prop. 8]. The result that the Hadamard product of a rational series and an algebraic series is algebraic (Proposition 6.6.12) is due to Schützenberger [**65**]. For further information on "closure properties" of formal series, see Exercise 6.71.

Our main result on noncommutative diagonals (Theorem 6.7.1) is a special case of a theorem of G. Jacob [**31**, Thm. 4]. We have closely followed Haiman [**35**] in our development of the theory of Section 6.7. For further information on Example 6.7.3, see Exercise 6.74.

References

6.1. D. Aldous, Triangulating the circle at random, *Amer. Math. Monthly* **101** (1994), 223–233.
6.2. R. Alter, Some remarks and results on Catalan numbers, in *Proc. Second Louisiana Conference on Combinatorics, Graph Theory and Computing (1971)* (R. C. Mullin *et al.*, eds.), Congressum Numeratium III, Utilitas Mathematica Publishing Co., Winnipeg, 1971, pp. 109–132.
6.3. F. Bergeron and C. Reutenauer, Combinatorial resolution of systems of differential equations. III. A special class of differentially algebraic series, *Europ. J. Combinatorics* **11** (1990), 501–512.
6.4. F. Bergeron and U. Sattler, Constructible differentially finite algebraic series in several variables, *Theoret. Comput. Sci.* **144** (1995), 59–65.
6.5. J. Berstel and C. Reutenauer, *Rational Series and Their Languages*, Springer-Verlag, Berlin/Heidelberg/New York, 1988.
6.6. J. Bertrand, Calcul des probabilités. Solution d'un problème, *C. R. Acad. Sci. Paris* **105** (1887), 369.
6.7. J. P. M. Binet, Sur un problème traité autrefois par Euler et par Ségner, ayant pour objet la détermination du nombre des manières dont un polygone plan et rectiligne peut être partagé en triangles par ses diagonales, in *Extraits des Procès-Verbaux des Sciences de la Société Philomatique de Paris*, 1838, pp. 127–129.

6.8. J. P. M. Binet, Réflexions sur le problème de déterminer le nombre de manières dont une figure rectiligne peut être partagée en triangles au moyen de ses diagonales, *J. Math. Pures Appl.* **4** (1839), 79–91.

6.9. W. G. Brown, Historical note on a recurrent combinatorial problem, *Amer. Math. Monthly* **72** (1965), 973–977.

6.10. E. C. Catalan, Note sur une équation aux différences finies, *J. Math. Pures Appl. (1)* **3** (1838), 508–515; **4** (1839), 95–99.

6.11. A. Cayley, On the analytical form called trees, Part II, *Philos. Mag. (4)* **18** (1859), 374–378.

6.12. C. Chevalley, *Introduction to the Theory of Algebraic Functions of One Variable*, American Mathematical Society, Providence, 1951.

6.13. N. Chomsky and M. P. Schützenberger, The algebraic theory of context-free languages, in *Computer Programming and Formal Systems* (P. Braffort and D. Hirschberg, eds.), North-Holland, Amsterdam, 1963, pp. 118–161.

6.14. G. Chrystal, *Textbook of Algebra*, Part 2, Dover, New York, 1961.

6.15. J. Cofman, Catalan numbers for the classroom?, *Elemente der Mathematik* **52** (1997), 108–117.

6.16. P. M. Cohn, Algebra and language theory, *Bull. London Math. Soc.* **7** (1975), 1–29.

6.17. P. M. Cohn, Puiseux's theorem revisited, *J. Pure Appl. Algebra* **31** (1984), 1–4.

6.18. P. M. Cohn, Correction to "Puiseux's theorem revisited," *J. Pure Appl. Algebra* **52** (1988), 197–198.

6.19. L. Comtet, Calcul pratique des coefficients de Taylor d'une fonction algébrique, *Ens. Math.* **10** (1964), 267–270.

6.20. J. H. Conway and H. S. M. Coxeter, Triangulated polygons and frieze patterns, *Math. Gaz.* **57** (1973), 87–94, 175–183.

6.21. M. Eichler, *Introduction to the Theory of Algebraic Numbers and Functions*, Academic Press, New York/London, 1966.

6.22. G. Eisenstein, Allgemeine Auflösung der Gleichungen von den ersten vier Graden, *J. Reine Angew. Math.* **27** (1844), 81–83; reprinted in G. Eisenstein, *Mathematische Werke*, vol. 1., Chelsea, New York, 1975, pp. 7–9.

6.23. L. Euler, Summarium, *Novi Commentarii academiae scientiarum Petropolitanae* **7** (1758/59), 13–15. Reprinted in *Opera Omnia (1)* **26** (1953), xvi–xviii.

6.24. M. Fliess, Sur divers produits de séries formelles, *Bull. Soc. Math. France* **102** (1974), 181–191.

6.25. H. Furstenburg, Algebraic functions over finite fields, *J. Algebra* **7** (1967), 271–277.

6.26. N. von Fuss, Solutio quaestionis, quot modis polygonum n laterum in polygone m laterum per diagonales resolvi quest, *Novi Acta Petropolitanae* **9** (1795).

6.27. M. Gardner, Catalan numbers, Mathematical Games, *Scientific American* **234** (June, 1976), pp. 120–125, and bibliography on p. 132. Reprinted (with an Addendum) in *Time Travel and Other Mathematical Bewilderments*, W. H. Freeman, New York, 1988, Chapter 20 .

6.28. I. M. Gelfand, M. M. Kapranov, and A. V. Zelevinsky, *Discriminants, Resultants and Multidimensional Determinants*, Birkhäuser Boston, Cambridge, Massachusetts, 1994.

6.29. I. M. Gessel, A factorization for formal Laurent series and lattice path enumeration, *J. Combinatorial Theory (A)* **28** (1980), 321–337.

6.30. H. W. Gould, *Bell & Catalan Numbers*, revised ed., Combinatorial Research Institute, Morgantown, West Virginia, 1977.

6.31. G. Jacob, Sur un Théorème de Shamir, *Information and Control* **27** (1975), 218–261.

6.32. L. Habsieger, M. Kazarian, and S. Lando, On the second number of Plutarch, *American Math. Monthly* **105** (1998), 446.

6.33. B. Haible, *D-finite Potenzreihen in mehreren Variablen*, Diplomarbeit, Universität Karlsruhe, June, 1989.

6.34. B. Haible and M. Stoll, D-finite power series and the diagonal theorem, preprint dated 13 October 1993.
6.35. M. Haiman, Non-commutative rational power series and algebraic generating functions, *Europ. J. Combinatorics* **14** (1993), 335–339.
6.36. M. L. J. Hautus and D. A. Klarner, The diagonal of a double power series, *Duke Math. J.* **38** (1971), 229–235.
6.37. D. L. Hilliker, An algorithm for computing the values of the ramification index in the Puiseux series expansion of an algebraic function, *Pacific J. Math.* **118** (1985), 427–435.
6.38. P. J. Hilton and J. Pedersen, Catalan numbers, their generalization, and their uses, *Math. Intelligencer* **13** (Spring 1991), 64–75.
6.39. P. J. Hilton and J. Pedersen, Catalan-Zahlen und Wege in einem ganzzahligen Gitter, *Elemente der Math.* **48** (1993), 45–60.
6.40. R. Jungen, Sur la séries de Taylor n'ayant que des singularités algébrico-logarithmiques sur leur cercle de convergence, *Comment. Math. Helv.* **3** (1931), 266–306.
6.41. J. Koker, N. J. Kuenzi, A. Oktaç, R. Leibowitz, and L. Carmony, An investigation of the sequence of Catalan numbers with activities for prospective teachers, *School Science and Mathematics* **98** (1998), 4–11.
6.42. M. Kuchinski, *Catalan structures and correspondences*, M.Sc. thesis, Department of Mathematics, West Virginia University, 1977.
6.43. W. Kuich and A. Salomaa, *Semirings, Automata, Languages*, Springer-Verlag, Berlin/Heidelberg/New York, 1986.
6.44. G. Lamé, Extrait d'une lettre de M. Lamé à M. Liouville sur cette question: Un polygone convexe étant donné, de combien de manières peut-on le partager en triangles au moyen de diagonales?, *J. Math. Pures Appl. (1)* **3** (1838), 505–507.
6.45. S. Lang, *Algebra*, Addison-Wesley, Reading, Massachusetts, 1965.
6.46. C. W. Lee, The associahedron and triangulations of the n-gon, *Europ. J. Combinatorics* **10** (1989), 551–560.
6.47. S. Lefschetz, *Algebraic Geometry*, Princeton University Press, Princeton, New Jersey, 1953.
6.48. L. Lipshitz, D-finite power series, *J. Algebra* **122** (1989), 353–373.
6.49. L. Lipshitz and A. J. van der Poorten, Rational functions, diagonals, automata, and arithmetic, in *Number Theory* (R. A. Mollin, ed.), Gruyter, Berlin/New York, 1990, pp. 339–358.
6.50. Luo Jian-Jin, Catalan numbers in the history of mathematics in China, in *Combinatorics and Graph Theory*, Proc. Spring School and International Conference on Combinatorics, Hefei, 6–27 April 1992 (H. P. Yap *et al.*, eds.), World Scientific, 1992, pp. 68–70.
6.51. S. J. Patterson, Eisenstein and the quintic equation, *Historia Mathematica* **17** (1990), 132–140.
6.52. M. Petkovsek, H. Wilf, and D. Zeilberger, $A = B$, A K Peters Ltd., Wellesley, Massachusetts, 1996.
6.53. G. Pólya, Sur les séries entières dont la somme est une fonction algébrique, *Ens. Math.* **22** (1921–1922), 38–47; *Collected Papers*, vol. I, MIT Press, Cambridge, Massachusetts, 1974, pp. 165–174.
6.54. V. Puiseux, Recherches sur les fonctions algébriques, *J. Math. Pures Appl. (1)* **15** (1850), 365–480.
6.55. C. Reutenauer, A survey of noncommutative rational series, in *Formal Power Series and Algebraic Combinatorics (New Brunswick, NJ, 1994)*, DIMACS Series Discrete Math. Theoret. Comput. Sci. **24**, American Mathematical Society, Providence, 1996, pp. 159–169.
6.56. O. Rodrigues, Sur le nombre de manières de décomposer un polygone en triangles au moyen de diagonales, *J. Math. Pures Appl.* **3** (1838), 547–548.

6.57. O. Rodriques, Sur le nombre de manières d'effectuer un produit de n facteurs, *J. Math. Pures Appl.* **3** (1838), 549.
6.58. D. G. Rogers, A Schröder triangle: Three combinatorial problems, in *Combinatorial Mathematics V: Proceedings of the Fifth Australian Conference*, Lecture Notes in Math. **622**, Springer-Verlag, Berlin/Heidelberg/New York, 1977, pp. 175–196.
6.59. L. A. Rubel, Some research problems about algebraic differential equations, *Trans. Amer. Math. Soc.* **280** (1983), 43–52.
6.60. L. A. Rubel, Some research problems about algebraic differential equations. II, *Illinois J. Math.* **36** (1992), 659–680.
6.61. A. Salomaa, *Formal Languages*, Academic Press, New York/London, 1973.
6.62. A. Salomaa, *Computation and Automata*, Encyclopedia of Mathematics and Its Applications **25** (G.-C. Rota, ed.), Cambridge University Press, Cambridge, 1985.
6.63. A. Salomaa and M. Soittola, *Automata-Theoretic Aspects of Formal Power Series*, Springer-Verlag, Berlin/Heidelberg/New York, 1978.
6.64. M. P. Schützenberger, *Un Problème de la Théorie des Automates*, Séminaire Dubriel-Pisot, 13e année (1959–1960), no. 3, Inst. H. Poincaré, Paris, 1960.
6.65. M. P. Schützenberger, On a theorem of R. Jungen, *Proc. Amer. Math. Soc.* **13** (1962), 885–890.
6.66. M. P. Schützenberger, On the definition of a family of automata, *Information and Control* **4** (1961), 245–270.
6.67. A. de Segner, Enumeratio modorum, quibus figurae planae rectilineae per diagonales dividuntur in triangula, *Novi Commentarii Academiae Scientiarum Petropolitanae* **7** (1758/59), 203–209.
6.68. E. Shamir, Algebraic, rational, and context-free power series in noncommuting variables, in *Algebraic Theory of Machines, Languages, and Semigroups* (M. Arbib, ed.), Academic Press, New York, 1968, pp. 329–341.
6.69. D. D. Sleator, R. E. Tarjan, and W. P. Thurston, Rotation distance, triangulations and hyperbolic geometry, *J. Amer. Math. Soc.* **1** (1988), 647–681.
6.70. R. Stanley, Differentiably finite power series, *Europ. J. Combinatorics* **1** (1980), 175–188.
6.71. R. Stanley, Hipparchus, Plutarch, Schröder and Hough, *American Math. Monthly* **104** (1997), 344–350.
6.72. G. Viennot, Enumerative combinatorics and algebraic languages, in *Fundamentals of computation theory (Cottbus, 1985)*, Lecture Notes in Computer Science **199**, Springer-Verlag, Berlin/Heidelberg/New York, 1985, pp. 450–464.
6.73. I. Vun and P. Belcher, Catalan numbers, *Math. Spectrum* **30** (1997/98), 3–5.
6.74. R. J. Walker, *Algebraic Curves*, Princeton University Press, Princeton, New Jersey, 1950. Reprinted by Dover, New York, 1962.
6.75. D. Zeilberger, A holonomic systems approach to special function identities, *J. Comput. Appl. Math.* **32** (1990), 321–368.

Exercises

6.1. [2+] Give a formal proof (no complex analysis, etc.) that e^x is not algebraic, i.e.,

$$\sum_{n\geq 0} \frac{x^n}{n!} \notin \mathbb{C}_{\mathrm{alg}}[[x]].$$

6.2. **a.** [3−] Suppose that $F(x) \in \mathbb{Q}[[x]]$ is algebraic. Show that there is an integer $m \geq 1$ such that $F(mx) \in \mathbb{Z}[[x]]$.

b. [1] Deduce that e^x is not algebraic.

6.3. [3+] Show that $y_1 = \sum_{n\geq 0} \binom{3n}{n,n,n} x^n$ and $y_2 = \sum_{n\geq 0} \binom{2n}{n}^2 x^n$ are not algebraic (over a field of characteristic zero). What about $\sum_{n\geq 0} \binom{2n}{n}^3 x^n$?

6.4. [3−] Show that Puiseux's theorem (Theorem 6.1.5) fails in characteristic $p > 0$.

6.5. [2] Let char $K = 0$ and

$$P(y) = F_d(x)y^d + \cdots + F_0(x) \in K[[x]][y],$$

with $F_d(0) \neq 0$. Suppose that $P(y)$ is irreducible, and let $c_1, \ldots, c_r$ be given by Corollary 6.1.7. Show that disc $P(y)$ is divisible by x^{d-r}.

6.6. [2+] Show that $\sum_{n\geq 0} f(n)x^n \in K[[x]]$ is algebraic of degree d if and only if $\sum_{n\geq 0}(\Delta^n f(0))x^n$ is algebraic of degree d. (See Section 1.4 for the definition of $\Delta^n f(0)$.)

6.7. [2−] If $u \in K[[x]]$ is algebraic with $u(0) = 0$ and $u'(0) \neq 0$, then is the compositional inverse $u^{\langle -1\rangle}$ algebraic?

6.8. **a.** [3−] Verify equation (6.7), i.e., show that

$$\operatorname{disc}(ay^d + by + c) = (-1)^{\binom{d}{2}}a^{d-2}[d^d ac^{d-1} + (-1)^{d-1}(d-1)^{d-1}b^d].$$

b. [3−] Find $\operatorname{disc}\left(\frac{x^n}{n!} + \frac{x^{n-1}}{(n-1)!} + \cdots + x + 1\right)$.

6.9. **a.** [2+] Let $F(x, y) = \sum_{m,n\geq 0} f(m, n)x^m y^n$, $G(x, y) = \sum_{m,n\geq 0} g(m, n)x^m y^n \in K[[x, y]]$, where K is any field. The Hadamard product is defined analogously to the univariate case to be

$$F * G = \sum_{m,n\geq 0} f(m, n)g(m, n)x^m y^n.$$

Show that if F and G are rational, then $F * G$ is algebraic (over $K(x, y)$).

b. [3−] For $k \geq 2$ define the power series

$$F_k(x_1, \ldots, x_k) = \sum_{n_1,\ldots,n_k\geq 0} \binom{n_1+n_2}{n_1}\binom{n_2+n_3}{n_2}\cdots\binom{n_k+n_1}{n_k} x_1^{n_1}\cdots x_k^{n_k}.$$

Show that F_k is algebraic (over $K(x_1, \ldots, x_k)$).

c. [2] Compute $F_2(x, y) = \sum_{m,n\geq 0} \binom{m+n}{m}^2 x^m y^n$ and $F_3(x, y, z)$ explicitly.

6.10. [2] Let $P(q) \in K[q, q^{-1}]$ be a Laurent polynomial over K, and fix an integer m. Define $f(n) = [q^m]P(q)^n$ for all $n \geq 0$. Show that $y = \sum_{n\geq 0} f(n)x^n$ is algebraic.

6.11. **a.** [3] Let K be a field of characteristic $p > 0$. Let $F, G \in K_{\mathrm{alg}}[[x_1, \ldots, x_k]]$, i.e., F and G are algebraic power series over the field $K(x_1, \ldots, x_k)$. Show that the Hadamard product $F * G$ (defined for $k = 2$ in Exercise 6.9(a) and extended in the obvious way to arbitrary k) is also algebraic.

b. [2+] Deduce that if $F \in K_{\mathrm{alg}}[[x_1, \ldots, x_k]]$ (with char $K = p > 0$), then diag F is algebraic.

6.12. [2+] Given power series $F(x) = F(x_1, \ldots, x_m) \in K[[x_1, \ldots, x_m]]$ and $G(y) = G(y_1, \ldots, y_n) \in K[[y_1, \ldots, y_n]]$, let F_k denote the part of F that is homogeneous of degree k, and similarly G_k, so $F = \sum F_k$ and $G = \sum G_k$. Define

$$(F \heartsuit G)(x, y) = \sum_{k\geq 0} F_k(x)G_k(y) \in K[[x, y]],$$

the "heartamard product" of F and G. Show that if F and G are rational, then so is $F \heartsuit G$. Moreover, if F is rational and G is algebraic, then $F \heartsuit G$ is algebraic.

6.13. **a.** [3−] Let $k \in \mathbb{P}$, and define $\eta = \sum_{n \geq 0} \binom{kn}{n} x^n$. Example 6.2.7 shows that η is a root of the polynomial

$$P(y) = k^k x y^k - (y-1)[(k-1)y+1]^{k-1}.$$

Find (as fractional series) the other $k-1$ roots of the polynomial $P(y)$. Deduce that $P(y)$ is irreducible (as a polynomial over $\mathbb{C}(x)$).

b. [3−] Find the discriminant of $P(y)$.

6.14. [3−] Define $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Q}$ by

$$f(i-1, j) - 2f(i, j) + f(i+1, j-1) = 0 \tag{6.53}$$

for all $(i, j) \in \mathbb{N} \times \mathbb{N} - \{(0, 0)\}$, with the initial conditions $f(0, 0) = 1$ and $f(i, j) = 0$ if $i < 0$ or $j < 0$. Thus $f(i, 0) = 2^{-i}$, $f(0, 1) = \frac{1}{4}$, $f(1, 1) = \frac{1}{4}$, etc. Find the generating function $F(x, y) = \sum_{i,j \geq 0} f(i, j) x^i y^j$.

6.15. [2+] Let $f, g, h \in K[[x]]$ with $h(0) = 0$. Find a polynomial $P(f, g, h, x)$ so that

$$\operatorname{diag} \frac{1}{1 - sf(st) - tg(st) - h(st)} = \frac{1}{\sqrt{P}},$$

where diag is in the variable x.

6.16. [5−] Let $f(n)$ be the number of paths from $(0, 0)$ to (n, n) using the steps $(1, 0)$, $(0, 1)$, and $(1, 1)$; and let $g(n)$ be the number of paths from $(0, 0)$ to (n, n) using any elements of $\mathbb{N}^2 - \{(0, 0)\}$ as steps. It is immediate from equations (6.27) and (6.30) that $g(n) = 2^{n-1} f(n)$, $n > 0$. Is there a combinatorial proof?

6.17. **a.** [2+] Let S be a subset of $\mathbb{N} \times \mathbb{N} - \{(0, 0)\}$ such that (i) every element of S has the form (n, n), $(n+1, n)$, or $(n, n+1)$, and (ii) $(n, n+1) \in S$ if and only if $(n+1, n) \in S$. Let $g(n)$ be the number of paths from $(0, 0)$ to (n, n) using steps from S. Let $h(n)$ be the number of such paths that never go above the line $y = x$. (Let $g(0) = h(0) = 1$.) Define $G(x) = \sum_{n \geq 0} g(n) x^n$, $H(x) = \sum_{n \geq 0} h(n) x^n$, and $K(x) = \sum_{(n,n) \in S} x^n$. Show that

$$H(x) = \frac{2}{1 - K(x) + G(x)^{-1}}.$$

b. [2−] Compute $H(x)$ explicitly when $S = \{(0, 1), (1, 0), (1, 1)\}$ and deduce that in this case $h(n)$ is the Schröder number r_n, thus confirming Exercise 6.39(j).

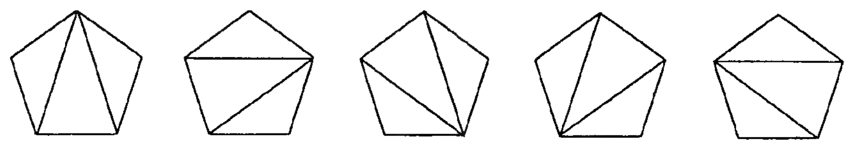
c. [3−] Give a *combinatorial* proof that when $S = \{(0, 1), (1, 0), (1, 1)\}$ and $n \geq 2$, then $h(n)$ is twice the number of ways to dissect a convex $(n+2)$-gon with any number of diagonals that don't intersect in their interiors.

6.18. [3] Let S be a subset of $\mathbb{N} \times \mathbb{N} - \{(0, 0)\}$ such that $\sum_{(m,n) \in S} x^m y^n$ is rational, e.g., S is finite or cofinite. Let $f(n)$ be the number of lattice paths from $(0, 0)$ to (n, n) with steps from S that never go above the line $y = x$. Show that $\sum_{n \geq 0} f(n) x^n$ is algebraic.

6.19. [1]–[3+] Show that the Catalan numbers $C_n = \frac{1}{n+1}\binom{2n}{n}$ count the number of elements of the 66 sets S_i, (a) $\leq i \leq$ (nnn), given below. We illustrate the elements of each S_i for $n = 3$, hoping that these illustrations will make any

undefined terminology clear. (The terms used in (vv)–(yy) are defined in Chapter 7.) Ideally S_i and S_j should be proved to have the same cardinality by exhibiting a simple, elegant bijection $\phi_{ij} : S_i \to S_j$ (so 4290 bijections in all). In some cases the sets S_i and S_j will actually coincide, but their descriptions will differ.

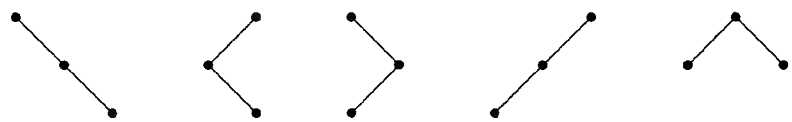
a. Triangulations of a convex $(n+2)$-gon into n triangles by $n-1$ diagonals that do not intersect in their interiors:



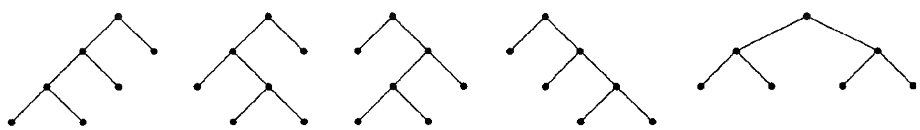
b. Binary parenthesizations of a string of $n+1$ letters:

$$(xx \cdot x)x \quad x(xx \cdot x) \quad (x \cdot xx)x \quad x(x \cdot xx) \quad xx \cdot xx$$

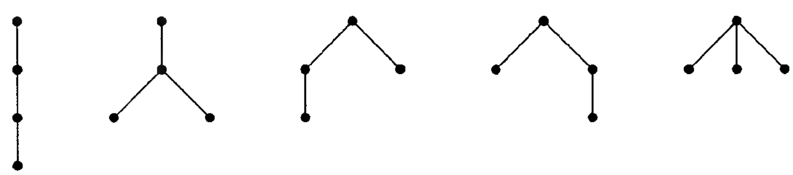
c. Binary trees with n vertices:



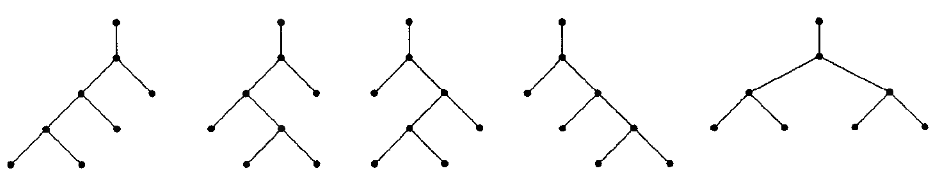
d. Plane binary trees with $2n+1$ vertices (or $n+1$ endpoints):



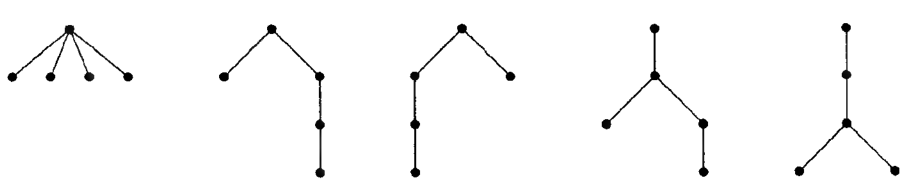
e. Plane trees with $n+1$ vertices:



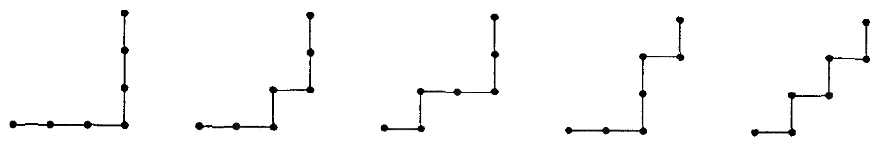
f. Planted (i.e., root has degree one) trivalent plane trees with $2n+2$ vertices:



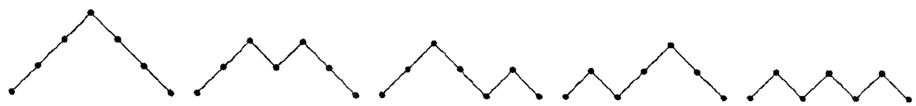
g. Plane trees with $n+2$ vertices such that the rightmost path of each subtree of the root has even length:



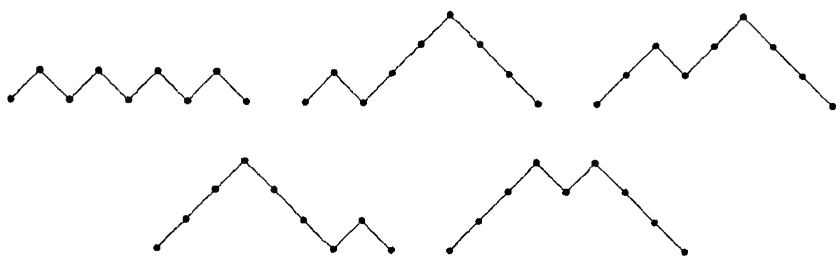
h. Lattice paths from $(0, 0)$ to (n, n) with steps $(0, 1)$ or $(1, 0)$, never rising above the line $y = x$:



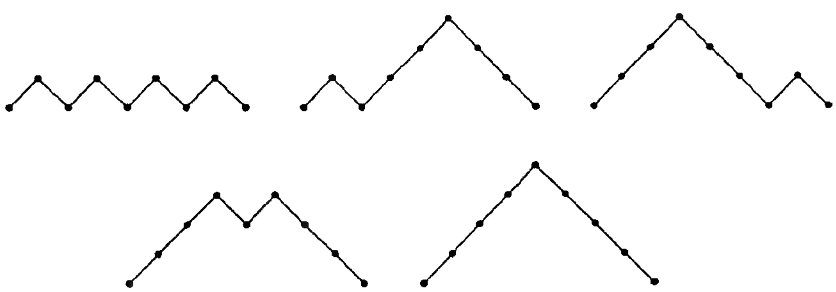
i. Dyck paths from $(0, 0)$ to $(2n, 0)$, i.e., lattice paths with steps $(1, 1)$ and $(1, -1)$, never falling below the x-axis:



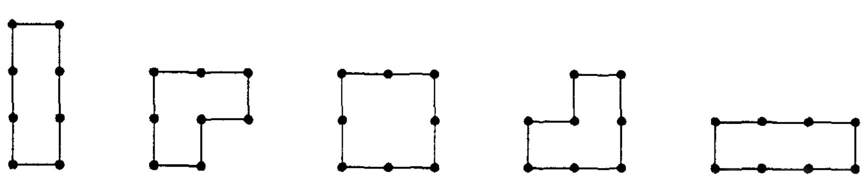
j. Dyck paths (as defined in (i)) from $(0, 0)$ to $(2n+2, 0)$ such that any maximal sequence of consecutive steps $(1, -1)$ ending on the x-axis has odd length:



k. Dyck paths (as defined in (i)) from $(0, 0)$ to $(2n + 2, 0)$ with no peaks at height two

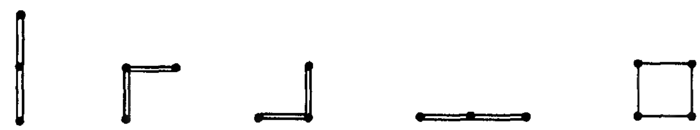


l. (Unordered) pairs of lattice paths with $n + 1$ steps each, starting at $(0, 0)$, using steps $(1, 0)$ or $(0, 1)$, ending at the same point, and only intersecting at the beginning and end:

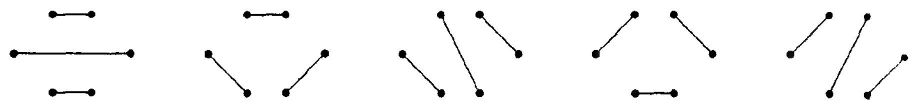


m. (Unordered) pairs of lattice paths with $n - 1$ steps each, starting at $(0, 0)$, using steps $(1, 0)$ or $(0, 1)$, ending at the same point, such that one path never

rises above the other path:



n. n nonintersecting chords joining $2n$ points on the circumference of a circle:



o. Ways of connecting $2n$ points in the plane lying on a horizontal line by n nonintersecting arcs, each arc connecting two of the points and lying above the points:

p. Ways of drawing in the plane $n+1$ points lying on a horizontal line L and n arcs connecting them such that (α) the arcs do not pass below L, (β) the graph thus formed is a tree, (γ) no two arcs intersect in their interiors (i.e., the arcs are noncrossing), and (δ) at every vertex, all the arcs exit in the same direction (left or right):

q. Ways of drawing in the plane $n+1$ points lying on a horizontal line L and n arcs connecting them such that (α) the arcs do not pass below L, (β) the graph thus formed is a tree, (γ) no arc (including its endpoints) lies strictly below another arc, and (δ) at every vertex, all the arcs exit in the same direction (left or right):

r. Sequences of n 1's and n -1's such that every partial sum is nonnegative (with -1 denoted simply as $-$ below):

$$111{-}{-}{-} \quad 11{-}1{-}{-} \quad 11{-}{-}1{-} \quad 1{-}11{-}{-} \quad 1{-}1{-}1{-}$$

s. Sequences $1 \le a_1 \le \cdots \le a_n$ of integers with $a_i \le i$:

111 112 113 122 123

t. Sequences $a_1 < a_2 < \cdots < a_{n-1}$ of integers satisfying $1 \le a_i \le 2i$:

12 13 14 23 24

u. Sequences $a_1, a_2, \ldots, a_n$ of integers such that $a_1 = 0$ and $0 \le a_{i+1} \le a_i + 1$:

000 001 010 011 012

v. Sequences $a_1, a_2, \ldots, a_{n-1}$ of integers such that $a_i \le 1$ and all partial sums are nonnegative:

$$0,0 \quad 0,1 \quad 1,-1 \quad 1,0 \quad 1,1$$

w. Sequences $a_1, a_2, \ldots, a_n$ of integers such that $a_i \ge -1$, all partial sums are nonnegative, and $a_1 + a_2 + \cdots + a_n = 0$:

$$0,0,0 \quad 0,1,-1 \quad 1,0,-1 \quad 1,-1,0 \quad 2,-1,-1$$

x. Sequences $a_1, a_2, \ldots, a_n$ of integers such that $0 \le a_i \le n-i$, and such that if $i < j$, $a_i > 0$, $a_j > 0$, and $a_{i+1} = a_{i+2} = \cdots = a_{j-1} = 0$, then $j - i > a_i - a_j$:

$$000 \quad 010 \quad 100 \quad 200 \quad 110$$

y. Sequences $a_1, a_2, \ldots, a_n$ of integers such that $i \le a_i \le n$ and such that if $i \le j \le a_i$, then $a_j \le a_i$:

$$123 \quad 133 \quad 223 \quad 323 \quad 333$$

z. Sequences $a_1, a_2, \ldots, a_n$ of integers such that $1 \le a_i \le i$ and such that if $a_i = j$, then $a_{i-r} \le j - r$ for $1 \le r \le j-1$:

$$111 \quad 112 \quad 113 \quad 121 \quad 123$$

aa. Equivalence classes B of words in the alphabet $[n-1]$ such that any three consecutive letters of any word in B are distinct, under the equivalence relation $uijv \sim ujiv$ for any words u, v and any $i, j \in [n-1]$ satisfying $|i-j| \ge 2$:

$$\{\emptyset\} \quad \{1\} \quad \{2\} \quad \{12\} \quad \{21\}$$

(For $n = 4$ a representative of each class is given by $\emptyset$, 1, 2, 3, 12, 21, 13, 23, 32, 123, 132, 213, 321, 2132.)

bb. Partitions $\lambda = (\lambda_1, \ldots, \lambda_{n-1})$ with $\lambda_1 \le n-1$ (so the diagram of λ is contained in an $(n-1) \times (n-1)$ square), such that if $\lambda' = (\lambda'_1, \lambda'_2, \ldots)$ denotes the conjugate partition to λ then $\lambda'_i \ge \lambda_i$ whenever $\lambda_i \ge i$:

$$(0,0) \quad (1,0) \quad (1,1) \quad (2,1) \quad (2,2)$$

cc. Permutations $a_1 a_2 \cdots a_{2n}$ of the multiset $\{1^2, 2^2, \ldots, n^2\}$ such that: (i) the first occurrences of $1, 2, \ldots, n$ appear in increasing order, and (ii) there is no subsequence of the form $\alpha\beta\alpha\beta$:

$$112233 \quad 112332 \quad 122331 \quad 123321 \quad 122133$$

dd. Permutations $a_1 a_2 \cdots a_{2n}$ of the set $[2n]$ such that: (i) $1, 3, \ldots, 2n-1$ appear in increasing order, (ii) $2, 4, \ldots, 2n$ appear in increasing order, and (iii) $2i - 1$ appears before $2i$, $1 \le i \le n$:

$$123456 \quad 123546 \quad 132456 \quad 132546 \quad 135246$$

ee. Permutations $a_1a_2\cdots a_n$ of $[n]$ with longest decreasing subsequence of length at most two (i.e., there does not exist $i < j < k$, $a_i > a_j > a_k$), called 321-*avoiding* permutations:

$$123 \quad 213 \quad 132 \quad 312 \quad 231$$

ff. Permutations $a_1a_2\cdots a_n$ of $[n]$ for which there does not exist $i < j < k$ and $a_j < a_k < a_i$ (called 312-*avoiding* permutations):

$$123 \quad 132 \quad 213 \quad 231 \quad 321$$

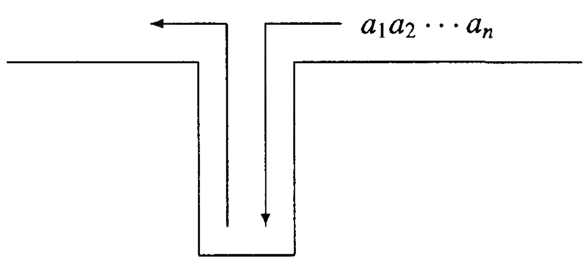
gg. Permutations w of $[2n]$ with n cycles of length two, such that the product $(1, 2, \dots, 2n) \cdot w$ has $n + 1$ cycles:

$$(1, 2, 3, 4, 5, 6)(1, 2)(3, 4)(5, 6) = (1)(2, 4, 6)(3)(5)$$
$$(1, 2, 3, 4, 5, 6)(1, 2)(3, 6)(4, 5) = (1)(2, 6)(3, 5)(4)$$
$$(1, 2, 3, 4, 5, 6)(1, 4)(2, 3)(5, 6) = (1, 3)(2)(4, 6)(5)$$
$$(1, 2, 3, 4, 5, 6)(1, 6)(2, 3)(4, 5) = (1, 3, 5)(2)(4)(6)$$
$$(1, 2, 3, 4, 5, 6)(1, 6)(2, 5)(3, 4) = (1, 5)(2, 4)(3)(6)$$

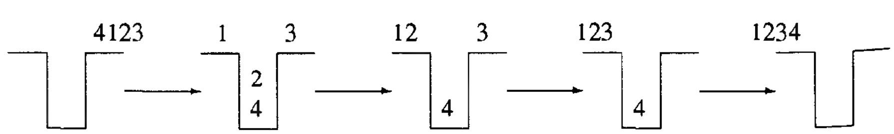
hh. Pairs (u, v) of permutations of $[n]$ such that u and v have a total of $n + 1$ cycles, and $uv = (1, 2, \dots, n)$:

$$(1)(2)(3) \cdot (1, 2, 3) \quad (1, 2, 3) \cdot (1)(2)(3) \quad (1, 2)(3) \cdot (1, 3)(2)$$
$$(1, 3)(2) \cdot (1)(2, 3) \quad (1)(2, 3) \cdot (1, 2)(3)$$

ii. Permutations $a_1a_2\cdots a_n$ of $[n]$ that can be put in increasing order on a single stack, defined recursively as follows: If $\emptyset$ is the empty sequence, then let $S(\emptyset) = \emptyset$. If $w = unv$ is a sequence of distinct integers with largest term n, then $S(w) = S(u)S(v)n$. A *stack-sortable* permutation w is one for which $S(w) = w$:

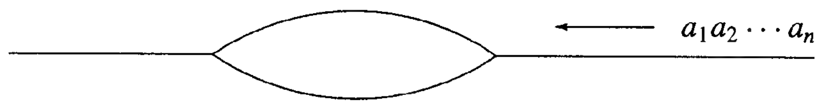


For example,



$$123 \quad 132 \quad 213 \quad 312 \quad 321$$

jj. Permutations $a_1a_2\cdots a_n$ of $[n]$ that can be put in increasing order on two parallel queues. Now the picture is



123 132 213 231 312

kk. Fixed-point-free involutions w of $[2n]$ such that if $i < j < k < l$ and $w(i) = k$, then $w(j) \neq l$ (in other words, 3412-avoiding fixed-point-free involutions):

(12)(34)(56) (14)(23)(56) (12)(36)(45) (16)(23)(45) (16)(25)(34)

ll. Cycles of length $2n+1$ in $\mathfrak{S}_{2n+1}$ with descent set $\{n\}$:

2371456 2571346 3471256 3671245 5671234

mm. Baxter permutations (as defined in Exercise 6.55) of $[2n]$ or of $[2n+1]$ that are reverse alternating (as defined at the end of Section 3.16) and whose inverses are reverse alternating:

132546 153426 354612 561324 563412

1325476 1327564 1534276 1735462 1756342

nn. Permutations w of $[n]$ such that if w has ℓ inversions then for all pairs of sequences $(a_1, a_2, \ldots, a_\ell), (b_1, b_2, \ldots, b_\ell) \in [n-1]^\ell$ satisfying

$$w = s_{a_1}s_{a_2}\cdots s_{a_\ell} = s_{b_1}s_{b_2}\cdots s_{b_\ell},$$

where s_j is the adjacent transposition $(j, j+1)$, we have that the ℓ-element *multisets* $\{a_1, a_2, \ldots, a_\ell\}$ and $\{b_1, b_2, \ldots, b_\ell\}$ are equal (thus, for example, $w = 321$ is not counted, since $w = s_1s_2s_1 = s_2s_1s_2$, and the multisets $\{1, 2, 1\}$ and $\{2, 1, 2\}$ are not equal):

123 132 213 231 312

oo. Permutations w of $[n]$ with the following property: Suppose that w has ℓ inversions, and let

$$R(w) = \left\{(a_1, \ldots, a_\ell) \in [n-1]^\ell : w = s_{a_1}s_{a_2}\cdots s_{a_\ell}\right\},$$

where s_j is as in (nn). Then

$$\sum_{(a_1,\ldots,a_\ell)\in R(w)} a_1a_2\cdots a_\ell = \ell!.$$

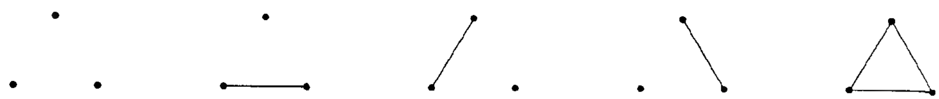
$$R(123) = \{\emptyset\}, \quad R(213) = \{(1)\}, \quad R(231) = \{(1,2)\}$$

$$R(312) = \{(2,1)\}, \quad R(321) = \{(1,2,1), (2,1,2)\}$$

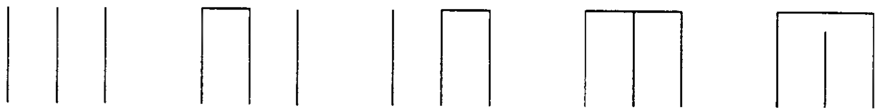
pp. Noncrossing partitions of $[n]$, i.e., partitions $\pi = \{B_1, \ldots, B_k\} \in \Pi_n$ such that if $a < b < c < d$ and $a, c \in B_i$ and $b, d \in B_j$, then $i = j$:

123 12–3 13–2 23–1 1–2–3

qq. Partitions $\{B_1, \ldots, B_k\}$ of $[n]$ such that if the numbers $1, 2, \ldots, n$ are arranged in order around a circle, then the convex hulls of the blocks $B_1, \ldots, B_k$ are pairwise disjoint:



rr. Noncrossing Murasaki diagrams with n vertical lines:



ss. Noncrossing partitions of some set $[k]$ with $n + 1$ blocks, such that any two elements of the same block differ by at least three:

1−2−3−4 14−2−3−5 15−2−3−4 25−1−3−4 16−25−3−4

tt. Noncrossing partitions of $[2n + 1]$ into $n + 1$ blocks, such that no block contains two consecutive integers:

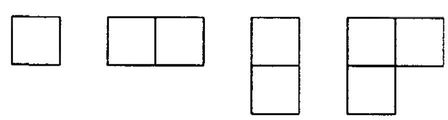
137−46−2−5 1357−2−4−6 157−24−3−6
17−246−3−5 17−26−35−4

uu. *Nonnesting partitions* of $[n]$, i.e., partitions of $[n]$ such that if a, e appear in a block B and b, d appear in a *different* block B' where $a < b < d < e$, then there is a $c \in B$ satisfying $b < c < d$:

123 12–3 13–2 23–1 1–2–3

(The unique partition of [4] that isn't nonnesting is 14–23.)

vv. Young diagrams that fit in the shape $(n - 1, n - 2, \ldots, 1)$:

$\emptyset$ 

ww. Standard Young tableaux of shape (n, n) (or equivalently, of shape $(n, n-1)$):

123 124 125 134 135
456 356 346 256 246

or

123 124 125 134 135
45 35 34 25 24

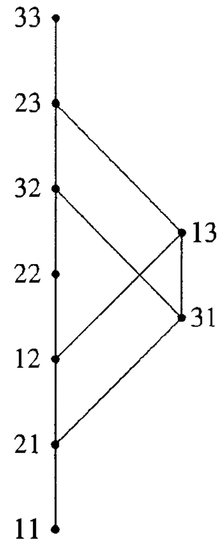


Figure 6-5. A poset with $C_4 = 14$ order ideals.

xx. Pairs (P, Q) of standard Young tableaux of the same shape, each with n squares and at most two rows:

$$(123, 123) \quad \left(\begin{matrix}12\\3\end{matrix}, \begin{matrix}12\\3\end{matrix}\right) \quad \left(\begin{matrix}12\\3\end{matrix}, \begin{matrix}13\\2\end{matrix}\right) \quad \left(\begin{matrix}13\\2\end{matrix}, \begin{matrix}12\\3\end{matrix}\right) \quad \left(\begin{matrix}13\\2\end{matrix}, \begin{matrix}13\\2\end{matrix}\right)$$

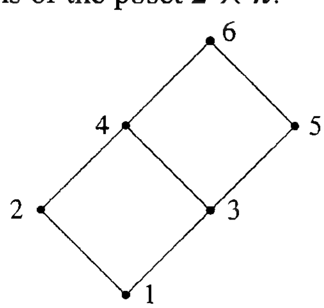
yy. Column-strict plane partitions of shape $(n-1, n-2, \ldots, 1)$, such that each entry in the i-th row is equal to $n-i$ or $n-i+1$:

$$\begin{matrix}33\\2\end{matrix} \quad \begin{matrix}33\\1\end{matrix} \quad \begin{matrix}32\\2\end{matrix} \quad \begin{matrix}32\\1\end{matrix} \quad \begin{matrix}22\\1\end{matrix}$$

zz. Convex subsets S of the poset $\mathbb{Z} \times \mathbb{Z}$, up to translation by a diagonal vector (m, m), such that if $(i, j) \in S$ then $0 < i - j < n$:

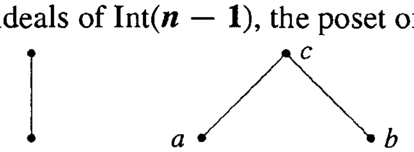
$$\emptyset \quad \{(1,0)\} \quad \{(2,0)\} \quad \{(1,0),(2,0)\} \quad \{(2,0),(2,1)\}$$

aaa. Linear extensions of the poset $\mathbf{2} \times \boldsymbol{n}$:



123456
123546
132456
132546
135246

bbb. Order ideals of Int($\boldsymbol{n} - \mathbf{1}$), the poset of intervals of the chain $\boldsymbol{n} - \mathbf{1}$:



2 Int(2)

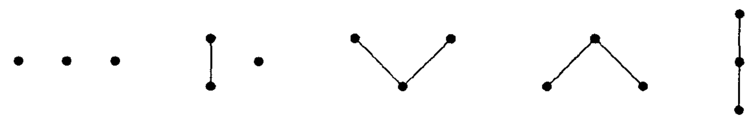
$\emptyset$, a, b, ab, abc

ccc. Order ideals of the poset A_n obtained from the poset $(\boldsymbol{n} - \mathbf{1}) \times (\boldsymbol{n} - \mathbf{1})$ by adding the relations $(i, j) < (j, i)$ if $i > j$ (see Figure 6-5 for the Hasse

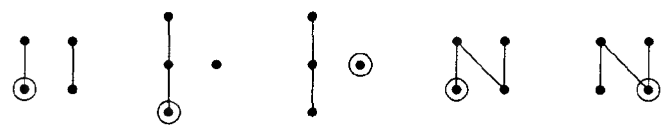
diagram of A_4):

$$\emptyset \quad \{11\} \quad \{11,21\} \quad \{11,21,12\} \quad \{11,21,12,22\}$$

ddd. Nonisomorphic n-element posets with no induced subposet isomorphic to $\mathbf{2}+\mathbf{2}$ or $\mathbf{3}+\mathbf{1}$:



eee. Nonisomorphic $(n+1)$-element posets that are a union of two chains and that are not a (nontrivial) ordinal sum, rooted at a minimal element:

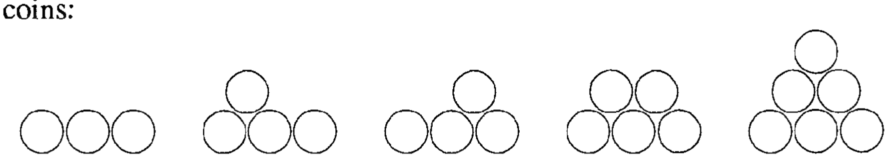


fff. Relations R on $[n]$ that are reflexive (iRi), symmetric ($iRj \Rightarrow jRi$), and such that if $1 \le i < j < k \le n$ and iRk, then iRj and jRk (in the example below we write ij for the pair (i,j), and we omit the pairs ii):

$$\emptyset \quad \{12,21\} \quad \{23,32\} \quad \{12,21,23,32\} \quad \{12,21,13,31,23,32\}$$

ggg. Joining some of the vertices of a convex $(n-1)$-gon by disjoint line segments, and circling a subset of the remaining vertices:

hhh. Ways to stack coins in the plane, the bottom row consisting of n consecutive coins:



iii. n-tuples $(a_1, a_2, \ldots, a_n)$ of integers $a_i \ge 2$ such that in the sequence $1a_1a_2\cdots a_n1$, each a_i divides the sum of its two neighbors:

14321 13521 13231 12531 12341

jjj. n-element multisets on $\mathbb{Z}/(n+1)\mathbb{Z}$ whose elements sum to 0:

000 013 022 112 233

kkk. n-element subsets S of $\mathbb{N}\times\mathbb{N}$ such that if $(i,j)\in S$ then $i \ge j$ and there is a lattice path from $(0,0)$ to (i,j) with steps $(0,1)$, $(1,0)$, and $(1,1)$ that lies entirely inside S:

$$\{(0,0),(1,0),(2,0)\} \quad \{(0,0),(1,0),(1,1)\} \quad \{(0,0),(1,0),(2,1)\}$$

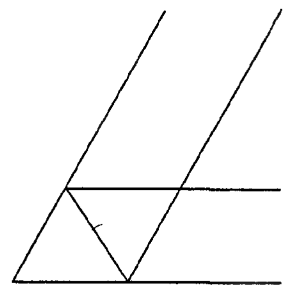
$$\{(0,0),(1,1),(2,1)\} \quad \{(0,0),(1,1),(2,2)\}$$

lll. Regions into which the cone $x_1 \ge x_2 \ge \cdots \ge x_n$ in $\mathbb{R}^n$ is divided by the hyperplanes $x_i - x_j = 1$, for $1 \le i < j \le n$ (the diagram below shows the

$$
\begin{array}{ccccccccccccccccccccccccccc}
1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 \\
& 1 & & 3 & & 2 & & 1 & & 5 & & 1 & & 2 & & 3 & & 1 & & 3 & & 2 & & 1 & & 5 & \\
& & 2 & & 5 & & 1 & & 4 & & 4 & & 1 & & 5 & & 2 & & 2 & & 5 & & 1 & & 4 & & \\
& & & 3 & & 2 & & 3 & & 3 & & 3 & & 2 & & 3 & & 3 & & 3 & & 2 & & 3 & & & \\
& & & & 1 & & 5 & & 2 & & 2 & & 5 & & 1 & & 4 & & 4 & & 1 & & 5 & & & & \\
& & & & & 2 & & 3 & & 1 & & 3 & & 2 & & 1 & & 5 & & 1 & & 2 & & & & & \\
& & & & & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & 1 & & & & & &
\end{array}
$$

Figure 6-6. The frieze pattern corresponding to the sequence $(1, 3, 2, 1, 5, 1, 2, 3)$.

situation for $n = 3$, intersected with the hyperplane $x_1 + x_2 + x_3 = 0$):



mmm. Positive integer sequences $a_1, a_2, \ldots, a_{n+2}$ for which there exists an integer array (necessarily with $n + 1$ rows)

$$
\begin{array}{cccccccccc}
1 \quad 1 & 1 & \cdots & 1 & 1 & 1 & \cdots & 1 & & 1 \\
a_1 \quad a_2 & a_3 & \cdots & a_{n+2} & a_1 & a_2 & \cdots & a_{n-1} & & \\
b_1 \quad b_2 & b_3 & \cdots & b_{n+2} & b_1 & \cdots & b_{n-2} & & & \\
& & \vdots & & & & & & & \\
r_1 & r_2 & r_3 & \cdots & r_{n+2} & r_1 & & & & \\
1 & 1 & 1 & \cdots & 1 & & & & &
\end{array}
\tag{6.54}
$$

such that any four neighboring entries in the configuration $\begin{smallmatrix} & r & \\ s & & t \\ & u & \end{smallmatrix}$ satisfy $st = ru + 1$ (an example of such an array for $(a_1, \ldots, a_8) = (1, 3, 2, 1, 5, 1, 2, 3)$ (necessarily unique) is given by Figure 6-6):

12213 22131 21312 13122 31221

nnn. n-tuples $(a_1, \ldots a_n)$ of positive integers such that the tridiagonal matrix

$$
\begin{bmatrix}
a_1 & 1 & 0 & 0 & \cdots & 0 & 0 \\
1 & a_2 & 1 & 0 & \cdots & 0 & 0 \\
0 & 1 & a_3 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & a_{n-1} & 1 \\
0 & 0 & 0 & 0 & \cdots & 1 & a_n
\end{bmatrix}
$$

is positive definite with determinant one:

131 122 221 213 312

6.20. **a.** [2+] Let m, n be integers satisfying $1 \le n < m$. Show by a simple bijection that the number of lattice paths from $(1, 0)$ to (m, n) with steps $(0, 1)$ and $(1, 0)$ that intersect the line $y = x$ in at least one point is equal to the number of lattice paths from $(0, 1)$ to (m, n) with steps $(0, 1)$ and $(1, 0)$.

b. [2–] Deduce that the number of lattice paths from $(0, 0)$ to (m, n) with steps $(1, 0)$ and $(0, 1)$ that intersect the line $y = x$ only at $(0, 0)$ is given by $\frac{m-n}{m+n}\binom{m+n}{n}$.

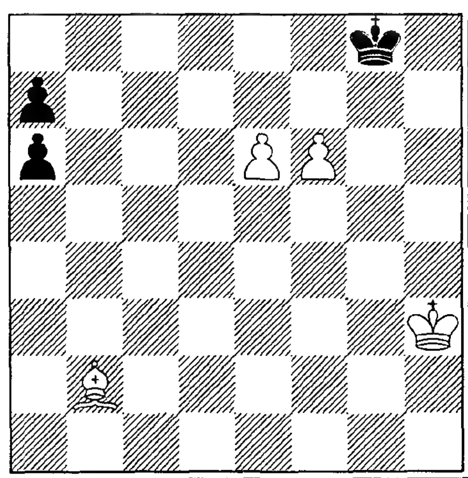
c. [1+] Show from (b) that the number of lattice paths from $(0, 0)$ to (n, n) with steps $(1, 0)$ and $(0, 1)$ that never rise above the line $y = x$ is given by the Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$. (This gives a direct combinatorial proof of interpretation (h) of C_n in Exercise 6.19.)

6.21. **a.** [2+] Let X_n be the set of all $\binom{2n}{n}$ lattice paths from $(0, 0)$ to (n, n) with steps $(0, 1)$ and $(1, 0)$. Define the *excedance* (also spelled "exceedance") of a path $P \in X_n$ to be the number of i such that at least one point (i, i') of P lies above the line $y = x$ (i.e., $i' > i$). Show that the number of paths in X_n with excedance j is independent of j.

b. [1] Deduce that the number of $P \in X_n$ that never rise above the line $y = x$ is given by the Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$ (a direct proof of interpretation (h) of C_n in Exercise 6.19). Compare with Example 5.3.11, which also gives a direct combinatorial interpretation of C_n when written in the form $\frac{1}{n+1}\binom{2n}{n}$ (as well as in the form $\frac{1}{2n+1}\binom{2n+1}{n}$).

6.22. [2+] Show (bijectively if possible) that the number of lattice paths from $(0, 0)$ to $(2n, 2n)$ with steps $(1, 0)$ and $(0, 1)$ that avoid the points $(2i - 1, 2i - 1)$, $1 \le i \le n$, is equal to the Catalan number C_{2n}.

6.23. [3–] Consider the following chess position:



Black is to make 19 consecutive moves, after which White checkmates Black in one move. Black may not move into check, and may not check White (except possibly on his last move). Black and White are *cooperating* to achieve the aim of checkmate. (In chess-problem parlance, this problem is called a *serieshelpmate in 19*.) How many different solutions are there?

6.24. [?] Explain the significance of the following sequence:

un, dos, tres, quatre, cinc, sis, set, vuit, nou, deu, ...

6.25. [2]–[5] Show that the Catalan number $C_n = \frac{1}{n+1}\binom{2n}{n}$ has the algebraic interpretations given below:

a. number of two-sided ideals of the algebra of all $(n-1) \times (n-1)$ upper triangular matrices over a field,

b. dimension of the space of invariants of SL(2, $\mathbb{C}$) acting on the $2n$-th tensor power $T^{2n}(V)$ of its "defining" two-dimensional representation V,

c. dimension of the irreducible representation of the symplectic group Sp($2(n-1)$, $\mathbb{C}$) (or Lie algebra $\mathfrak{sp}(2(n-1), \mathbb{C})$) with highest weight λ_{n-1}, the $(n-1)$-st fundamental weight,

d. dimension of the primitive intersection homology (say with real coefficients) of the toric variety associated with a (rationally embedded) n-dimensional cube,

e. the generic number of PGL(2, $\mathbb{C}$) equivalence classes of degree n rational maps with a fixed branch set,

f. number of translation conjugacy classes of degree $n+1$ monic polynomials in one complex variable, all of whose critical points are fixed,

g. dimension of the algebra (over a field K) with generators $\epsilon_1, \ldots, \epsilon_{n-1}$ and relations

$$\begin{aligned} \epsilon_i^2 &= \epsilon_i & & \\ \beta\epsilon_i\epsilon_j\epsilon_i &= \epsilon_i & & \text{if } |i-j| = 1 \\ \epsilon_i\epsilon_j &= \epsilon_j\epsilon_i & & \text{if } |i-j| \geq 2, \end{aligned}$$

where β is a nonzero element of K,

h. number of $\oplus$-sign types indexed by A_{n-1}^+ (the set of positive roots of the root system A_{n-1}).

i. Let the symmetric group $\mathfrak{S}_n$ act on the polynomial ring $A = \mathbb{C}[x_1, \ldots, x_n, y_1, \ldots, y_n]$ by $w \cdot f(x_1, \ldots, x_n, y_1, \ldots, y_n) = f(x_{w(1)}, \ldots, x_{w(n)}, y_{w(1)}, \ldots, y_{w(n)})$ for all $w \in \mathfrak{S}_n$. Let I be the ideal generated by all invariants of positive degree, i.e.,

$$I = \langle f \in A : w \cdot f = f \text{ for all } w \in \mathfrak{S}_n, \text{ and } f(0) = 0 \rangle.$$

Then (conjecturally) C_n is the dimension of the subspace of A/I affording the sign representation, i.e.,

$$C_n = \dim\{f \in A/I : w \cdot f = (\operatorname{sgn} w) f \text{ for all } w \in \mathfrak{S}_n\}.$$

6.26. a. [3−] Let D be a Young diagram of a partition λ, as defined in Section 1.3. Given a square s of D let t be the lowest square in the same column as s, and let u be the rightmost square in the same row as s. Let $f(s)$ be the number of paths from t to u that stay within D, and such that each step is one square to the north or one square to the east. Insert the number $f(s)$ in square s,

obtaining an array A. For instance, if $\lambda = (5, 4, 3, 3)$ then A is given by

16	7	2	1	1
6	3	1	1	
3	2	1		
1	1	1		

Let M be the largest square subarray (using consecutive rows and columns) of A containing the upper left-hand corner. Regard M as a matrix. For the above example we have

$$M = \begin{bmatrix} 16 & 7 & 2 \\ 6 & 3 & 1 \\ 3 & 2 & 1 \end{bmatrix}.$$

Show that $\det M = 1$.

b. [2] Find the unique sequence $a_0, a_1, \ldots$ of real numbers such that for all $n \geq 0$ we have

$$\det \begin{bmatrix} a_0 & a_1 & \cdots & a_n \\ a_1 & a_2 & \cdots & a_{n+1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_n & a_{n+1} & \cdots & a_{2n} \end{bmatrix} = \det \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ a_2 & a_3 & \cdots & a_{n+1} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_n & a_{n+1} & \cdots & a_{2n-1} \end{bmatrix} = 1.$$

(When $n = 0$ the second matrix is empty and by convention has determinant one.)

6.27. a. [3−] Let V_n be a real vector space with basis $x_0, x_1, \ldots, x_n$ and scalar product defined by $\langle x_i, x_j \rangle = C_{i+j}$, the $(i + j)$-th Catalan number. It follows from Exercise 6.26(b) that this scalar product is positive definite, and therefore V has an orthonormal basis. Is there an orthonormal basis for V_n whose elements are *integral* linear combinations of the x_i's ?

b. [3−] Same as (a), except now $\langle x_i, x_j \rangle = C_{i+j+1}$.

* **c.** [5−] Investigate the same question for the matrices M of Exercise 6.26(a) (so $\langle x_i, x_j \rangle = M_{ij}$) when λ is self-conjugate (so M is symmetric).

6.28. a. [3−] Suppose that real numbers $x_1, x_2, \ldots, x_d$ are chosen uniformly and independently from the interval [0, 1]. Show that the probability that the sequence $x_1, x_2, \ldots, x_d$ is convex (i.e., $x_i \leq \frac{1}{2}(x_{i-1} + x_{i+1})$ for $2 \leq i \leq d - 1$) is $C_{d-1}/(d-1)!^2$, where C_{d-1} denotes a Catalan number.

b. [3-] Let $\mathcal{C}_d$ denote the set of all points $(x_1, x_2, \ldots, x_d) \in \mathbb{R}^d$ such that $0 \leq x_i \leq 1$ and the sequence $x_1, x_2, \ldots, x_d$ is convex. It is easy to see that $\mathcal{C}_d$ is a d-dimensional convex polytope, called the *convexotope*. Show that

the vertices of $\mathcal{C}_d$ consist of the points

$$\left(1, \frac{j-1}{j}, \frac{j-2}{j}, \ldots, \frac{1}{j}, 0, 0, \ldots, 0, \frac{1}{k}, \frac{2}{k}, \ldots, 1\right) \tag{6.55}$$

(with at least one 0 coordinate), together with $(1, 1, \ldots, 1)$ (so $\binom{d+1}{2} + 1$ vertices in all). For instance, the vertices of $\mathcal{C}_3$ are $(0, 0, 0)$, $(0, 0, 1)$, $\left(0, \frac{1}{2}, 1\right)$, $(1, 0, 0)$, $\left(1, \frac{1}{2}, 0\right)$, $(1, 0, 1)$, $(1, 1, 1)$.

c. [3] Show that the Ehrhart quasi-polynomial $i(\mathcal{C}_d, n)$ of $\mathcal{C}_d$ (as defined in Section 4.6) is given by

$$\begin{aligned} y_d &:= \sum_{n \geq 0} i(\mathcal{C}_d, n) x^n \\ &= \frac{1}{1-x}\left(\sum_{r=1}^{d} \frac{1}{[1][r-1]!} * \frac{1}{[1][d-r]!}\right. \\ &\qquad \left. - \sum_{r=1}^{d-1} \frac{1}{[1][r-1]!} * \frac{1}{[1][d-1-r]!}\right), \end{aligned} \tag{6.56}$$

where $[i] = 1 - x^i$, $[i]! = [1][2]\cdots[i]$, and $*$ denotes Hadamard product. For instance,

$$y_1 = \frac{1}{(1-x)^2}$$

$$y_2 = \frac{1+x}{(1-x)^3}$$

$$y_3 = \frac{1+2x+3x^2}{(1-x)^3(1-x^2)}$$

$$y_4 = \frac{1+3x+9x^2+12x^3+11x^4+3x^5+x^6}{(1-x)^2(1-x^2)^2(1-x^3)}$$

$$y_5 = \frac{1+4x+14x^2+34x^3+63x^4+80x^5+87x^6+68x^7+42x^8+20x^9+7x^{10}}{(1-x)(1-x^2)^2(1-x^3)^2(1-x^4)}.$$

Is there a simpler formula than (6.56) for $i(\mathcal{C}_d, n)$ or y_d?

6.29. [3] Suppose that $n + 1$ points are chosen uniformly and independently from inside a square. Show that the probability that the points are in convex position (i.e., each point is a vertex of the convex hull of all the points) is $(C_n/n!)^2$.

6.30. [3−] Let f_n be the number of partial orderings of the set $[n]$ that contain no induced subposets isomorphic to $\mathbf{3} + \mathbf{1}$ or $\mathbf{2} + \mathbf{2}$. (This exercise is the labeled analogue of Exercise 6.19(ddd). As mentioned in the solution to this exercise, such posets are called *semiorders*.) Let $C(x) = 1 + x + 2x^2 + 5x^3 + \cdots$ be the generating function for Catalan numbers. Show that

$$\sum_{n \geq 0} f_n \frac{x^n}{n!} = C(1 - e^{-x}), \tag{6.57}$$

the composition of $C(x)$ with the series $1 - e^{-x} = x - \frac{1}{2}x^2 + \frac{1}{6}x^3 - \cdots$.

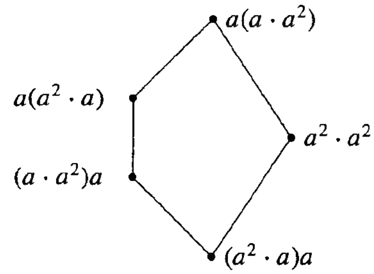


Figure 6-7. The Tamari lattice T_3.

6.31. **a.** [3−] Let $\mathcal{P}$ denote the convex hull in $\mathbb{R}^{d+1}$ of the origin together with all vectors $e_i - e_j$, where e_i is the i-th unit coordinate vector and $i < j$. Thus $\mathcal{P}$ is a d-dimensional convex polytope. Show that the relative volume of $\mathcal{P}$ (as defined in Section 4.6) is equal to $C_d/d!$, where C_d denotes a Catalan number.

b. [3] Let $i(\mathcal{P}, n)$ denote the Ehrhart polynomial of $\mathcal{P}$. Find a combinatorial interpretation of the coefficients of the i-Eulerian polynomial (in the terminology of Section 4.3)

$$(1-x)^{d+1} \sum_{n \geq 0} i(\mathcal{P}, n) x^n.$$

6.32. **a.** [3−] Define a partial order T_n on the set of all binary bracketings (parenthesizations) of a string of length $n+1$ as follows. We say that v covers u if u contains a subexpression $(xy)z$ (where x, y, z are bracketed strings) and v is obtained from u by replacing $(xy)z$ with $x(yz)$. For instance, $((a^2 \cdot a)a^2)(a^2 \cdot a^2)$ is covered by $((a \cdot a^2)a^2)(a^2 \cdot a^2)$, $(a^2(a \cdot a^2))(a^2 \cdot a^2)$, $((a^2 \cdot a)a^2)(a(a \cdot a^2))$, and $(a^2 \cdot a)(a^2(a^2 \cdot a^2))$. Figures 6-7 and 6-8 show the Hasse diagrams of T_3 and T_4. (In Figure 6-8, we have encoded the binary bracketing by a string of four $+$'s and four $-$'s, where a $+$ stands for a left parenthesis and a $-$ for the letter a, with the last a omitted.) Let U_n be the poset of all integer vectors $(a_1, a_2, \ldots, a_n)$ such that $i \leq a_i \leq n$ and such that if $i \leq j \leq a_i$ then $a_j \leq a_i$, ordered coordinatewise. Show that T_n and U_n are isomorphic posets.

b. [2] Deduce from (a) that T_n is a lattice (called the *Tamari lattice*).

6.33. Let C be a convex n-gon. Let $\mathcal{S}$ be the set of all sets of diagonals of C that do not intersect in the interior of C. Partially order the element of $\mathcal{S}$ by inclusion, and add a $\hat{1}$. Call the resulting poset A_n.

a. [3−] Show that A_n is a simplicial Eulerian lattice of rank $n-2$, as defined in Section 3.14.

b. [3] Show in fact that A_n is the lattice of faces of an $(n-3)$-dimensional convex polytope $\mathcal{Q}_n$.

c. [3−] Find the number $W_i = W_i(n)$ of elements of A_n of rank i. Equivalently, W_i is the number of ways to draw i diagonals of C that do not intersect in their interiors. Note that by Proposition 6.2.1, $W_i(n)$ is also the number of plane trees with $n+i$ vertices and $n-1$ endpoints such that no vertex has exactly one successor.

d. [3−] Define

$$\sum_{i=0}^{n-3} W_i (x-1)^{n-i-3} = \sum_{i=0}^{n-3} h_i x^{n-3-i}, \tag{6.58}$$

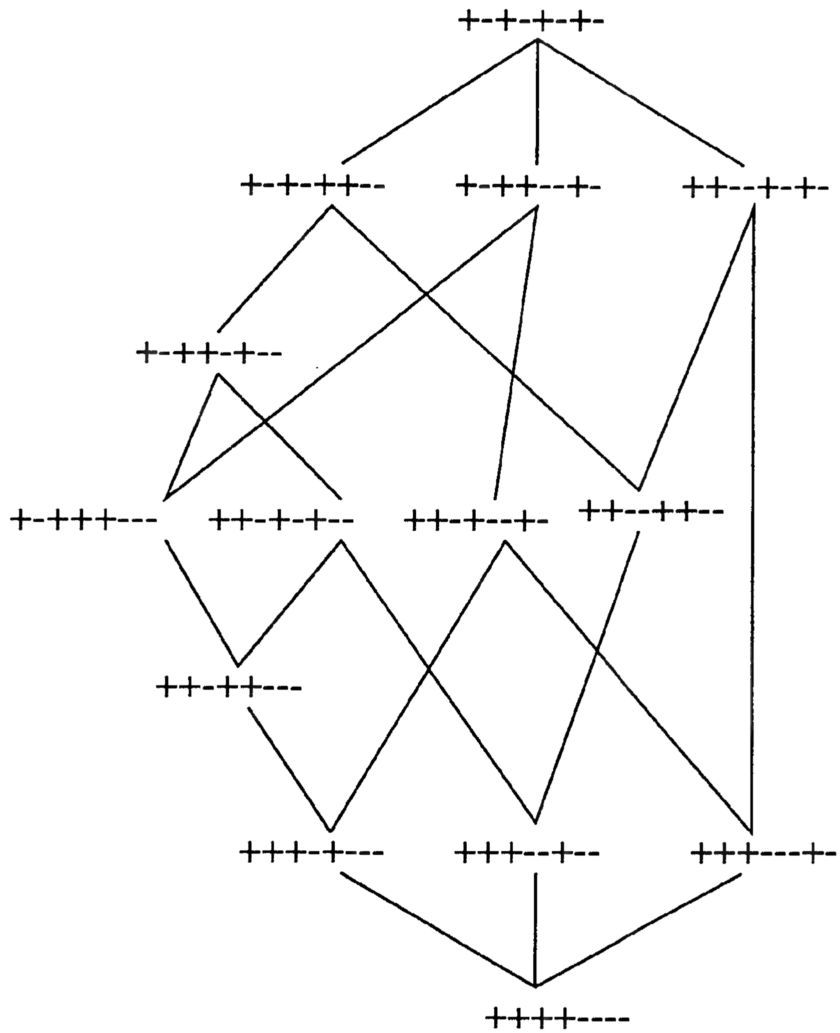


Figure 6-8. The Tamari lattice T_4.

as in equation (3.44). The vector $(h_0, \ldots, h_{n-3})$ is called the *h-vector* of A_n (or of the polytope $\mathcal{Q}_n$). Find an explicit formula for each h_i.

.34. There are many possible q-analogues of Catalan numbers. In (a) we give what is perhaps the most natural "combinatorial" q-analogue, while in (b) we give the most natural "explicit formula" q-analogue. In (c) we give an interesting extension of (b), while (d) and (e) are concerned with another special case of (c).

a. [2+] Let

$$C_n(q) = \sum_P q^{A(P)},$$

where the sum is over all lattice paths P from (0, 0) to (n, n) with steps (1, 0) and (0, 1), such that P never rises above the line $y = x$, and where $A(P)$ is the area under the path (and above the x-axis). Note that by Exercise 6.19(h), we have $C_n(1) = C_n$. (It is interesting to see what statistic corresponds to $A(P)$ for many of the other combinatorial interpretations of C_n given in Exercise 6.19.) For instance, $C_0(q) = C_1(q) = 1$, $C_2(q) = 1+q$, $C_3(q) = 1+q+2q^2+q^3$, $C_4(q) = 1+q+2q^2+3q^3+3q^4+3q^5+q^6$. Show

$x^{\alpha^4} = x_1x_2, x^{\alpha^5} = x_1x_4, x^{\alpha^6} = x_1, x^{\alpha^7} = x_3$. Then T is given by

$$\begin{matrix} 1\,1\,1\,3 \\ 2\,2\,3\,5 \\ 4\,4\,7 \\ 5 \\ 6 \end{matrix}$$

It is easy to see that the map $(\alpha^1, \alpha^2, \ldots) \mapsto T(\alpha^1, \alpha^2, \ldots)$ gives a bijection between ways of building up the term $x^{\lambda+\delta}$ from x^δ (according to the rules above) and SSYT of shape λ' and type μ, so the proof follows. $\square$

From the combinatorial definition of Schur functions it is clear that $s_\lambda(x_1, \ldots, x_n) = 0$ if $\ell(\lambda) > n$. Since by Proposition 7.8.2(b) we have $\dim \Lambda_n = \#\{\lambda \in \text{Par} : \ell(\lambda) \leq n\}$, it follows that the set $\{s_\lambda(x_1, \ldots, x_n) : \ell(\lambda) \leq n\}$ is a basis for Λ_n. (This also follows from a simple extension of the proof of Corollary 7.10.6.) We define on Λ_n a scalar product $\langle\, , \,\rangle_n$ by requiring that $\{s_\lambda(x_1, \ldots, x_n)\}$ is an orthonormal basis. If $f, g \in \Lambda$, then we write $\langle f, g\rangle_n$ as short for $\langle f(x_1, \ldots, x_n), g(x_1, \ldots, x_n)\rangle_n$. Thus

$$\langle f, g\rangle = \langle f, g\rangle_n,$$

provided that every monomial appearing in f involves at most n distinct variables, e.g., if $\deg f \leq n$.

7.15.2 Corollary. *If $f \in \Lambda_n$, $\ell(\lambda) \leq n$, and $\delta = (n-1, n-2, \ldots, 1, 0)$, then*

$$\langle f, s_\lambda\rangle_n = [x^{\lambda+\delta}]a_\delta f,$$

the coefficient of $x^{\lambda+\delta}$ in $a_\delta f$.

Proof. All functions will be in the variables $x_1, \ldots, x_n$. Let $f = \sum_{\ell(\lambda)\leq n} c_\lambda s_\lambda$. Then by Theorem 7.15.1 we have

$$a_\delta f = \sum_{\ell(\lambda)\leq n} c_\lambda a_{\lambda+\delta},$$

so

$$\langle f, s_\lambda\rangle_n = c_\lambda = [x^{\lambda+\delta}]a_\delta f. \qquad \square$$

For instance, we have

$$\left\langle a_\delta^{2k}, s_\lambda\right\rangle_n = [x^{\lambda+\delta}]a_\delta^{2k+1}, \tag{7.57}$$

for $\ell(\lambda) \leq n$. It is an interesting problem (not completely solved) to compute the numbers (7.57); for further information on the case $k = 1$, see Exercise 7.37.

d. [3−] Show that

$$c_n(0;q) = \frac{1+q}{1+q^n} c_n(q).$$

For instance, $c_0(0;q) = c_1(0;q) = 1$, $c_2(0;q) = 1+q$, $c_3(0;q) = 1+q+q^2+q^3+q^4$, $c_4(0;q) = 1+q+q^2+2q^3+2q^4+2q^5+2q^6+q^7+q^8+q^9$.

e. [3+] Show that the coefficients of $c_n(0;q)$ are *unimodal*, i.e., if $c_n(0;q) = \sum b_i q^i$, then for some j we have $b_0 \le b_1 \le \cdots \le b_j \ge b_{j+1} \ge b_{j+2} \ge \cdots$. (In fact, we can take $j = \lfloor \frac{1}{2} \deg c_n(0;q) \rfloor = \lfloor \frac{1}{2}(n-1)^2 \rfloor$.)

6.35. Let Q_n be the poset of direct-sum decompositions of an n-dimensional vector space V_n over the field $\mathbb{F}_q$, as defined in Example 5.5.2(b). Let $\bar{Q}_n$ denote Q_n with a $\hat{0}$ adjoined, and let $\mu_n(q) = \mu_{\bar{Q}_n}(\hat{0}, \hat{1})$. Hence by (5.74) we have

$$-\sum_{n\ge 1} \mu_n(q) \frac{x^n}{q^{\binom{n}{2}} (\boldsymbol{n})!} = \log \sum_{n\ge 0} \frac{x^n}{q^{\binom{n}{2}} (\boldsymbol{n})!}.$$

a. [3−] Show that

$$\mu_n(q) = \frac{1}{n}(-1)^n (q-1)(q^2-1)\cdots(q^{n-1}-1) P_n(q),$$

where $P_n(q)$ is a polynomial in q of degree $\binom{n}{2}$ with nonnegative integral coefficients, satisfying $P_n(1) = \binom{2n-1}{n}$. For instance,

$$\begin{aligned}
P_1(q) &= 1\\
P_2(q) &= 2+q\\
P_3(q) &= 3+3q+3q^2+q^3\\
P_4(q) &= (2+2q^2+q^3)(2+2q+2q^2+q^3).
\end{aligned}$$

b. [3−] Show that

$$\exp \sum_{n\ge 1} q^{\binom{n}{2}} P_n(1/q) \frac{x^n}{n} = \sum_{n\ge 1} q^{\binom{n}{2}} C_n(1/q) x^n,$$

where $C_n(q)$ is the q-Catalan polynomial defined in Exercise 6.34(a).

6.36. a. [2+] The *Narayana numbers* $N(n,k)$ are defined by

$$N(n,k) = \frac{1}{n}\binom{n}{k}\binom{n}{k-1}.$$

Let X_{nk} be the set of all sequences $w = w_1 w_2 \cdots w_{2n}$ of n 1's and n −1's with all partial sums nonnegative such that

$$k = \#\{j : w_j = 1, w_{j+1} = -1\}.$$

Give a combinatorial proof that $N(n,k) = \#X_{nk}$. Hence by Exercise 6.19(r), there follows

$$\sum_{k=1}^{n} N(n,k) = C_n.$$

(It is interesting to find for each of the combinatorial interpretations of C_n given by Exercise 6.19 a corresponding decomposition into subsets counted by Narayana numbers.)

b. [2+] Let $F(x,t) = \sum_{n\geq 1}\sum_{k\geq 1} N(n,k)x^n t^k$. Using the combinatorial interpretation of $N(n,k)$ given in (a), show that

$$xF^2 + (xt + x - 1)F + xt = 0, \tag{6.60}$$

so

$$F(x,t) = \frac{1 - x - xt - \sqrt{(1-x-xt)^2 - 4x^2t}}{2x}.$$

6.37. [2+] The *Motzkin numbers* M_n are defined by

$$\begin{aligned}\sum_{n\geq 0} M_n x^n &= \frac{1 - x - \sqrt{1 - 2x - 3x^2}}{2x^2}\\ &= 1 + x + 2x^2 + 4x^3 + 9x^4 + 21x^5 + 51x^6 + 127x^7\\ &\quad + 323x^8 + 835x^9 + 2188x^{10} + \cdots.\end{aligned}$$

Show that $M_n = \Delta^n C_1$ and $C_n = \Delta^{2n} M_0$, where C_n denotes a Catalan number.

6.38. [3−] Show that the Motzkin number M_n has the following combinatorial interpretations. (See Exercise 6.46(b) for an additional interpretation.)

a. Number of ways of drawing any number of nonintersecting chords among n points on a circle.

b. Number of walks on $\mathbb{N}$ with n steps, with steps -1, 0, or 1, starting and ending at 0.

c. Number of lattice paths from $(0,0)$ to (n,n), with steps $(0,2)$, $(2,0)$, and $(1,1)$, never rising above the line $y = x$.

d. Number of paths from $(0,0)$ to (n,n) with steps $(1,0)$, $(1,1)$, and $(1,-1)$, never going below the x-axis. Such paths are called *Motzkin paths*.

e. Number of pairs $1 \leq a_1 < \cdots < a_k \leq n$ and $1 \leq b_1 < \cdots < b_k \leq n$ of integer sequences such that $a_i \leq b_i$ and every integer in the set $[n]$ appears at least once among the a_i's and b_i's.

f. Number of ballot sequences (as defined in Corollary 6.2.3(ii)) $(a_1, \ldots, a_{2n+2})$ such that we never have $(a_{i-1}, a_i, a_{i+1}) = (1, -1, 1)$.

g. Number of plane trees with $n/2$ edges, allowing "half edges" that have no successors and count as half an edge.

h. Number of plane trees with $n+1$ edges in which no vertex, the root excepted, has exactly one successor.

i. Number of plane trees with n edges in which every vertex has at most two successors.

j. Number of binary trees with $n-1$ edges such that no two consecutive edges slant to the right.

k. Number of plane trees with $n+1$ vertices such that every vertex of odd height (with the root having height 0) has at most one successor.

l. Number of noncrossing partitions $\pi = \{B_1, \ldots, B_k\}$ of $[n]$ (as defined in Exercise 3.68) such that if $B_i = \{b\}$ and $a < b < c$, then a and c appear in different blocks of π.

m. Number of noncrossing partitions π of $[n+1]$ such that no block of π contains two consecutive integers.

6.39. [3−] The Schröder numbers r_n and s_n were defined in Section 6.2. Show that they have the following combinatorial interpretations.

a. s_{n-1} is the total number of bracketings (parenthesizations) of a string of n letters.

b. s_{n-1} is the number of plane trees with no vertex of degree one and with n endpoints.

c. r_{n-1} is the number of plane trees with n vertices and with each endpoint colored red or blue.

d. s_n is the number of binary trees with n vertices and with each right edge colored either red or blue.

e. s_n is the number of lattice paths in the (x, y)-plane from $(0, 0)$ to the x-axis using steps $(1, k)$, where $k \in \mathbb{P}$ or $k = -1$, never passing below the x-axis, and with n steps of the form $(1, -1)$.

f. s_n is the number of lattice paths in the (x, y)-plane from $(0, 0)$ to (n, n) using steps $(k, 0)$ or $(0, 1)$ with $k \in \mathbb{P}$, and never passing above the line $y = x$.

g. r_{n-1} is the number of parallelogram polynominoes (defined in the solution to Exercise 6.19(l)) of perimeter $2n$ with each column colored either black or white.

h. s_n is the number of ways to draw any number of diagonals of a convex $(n+2)$-gon that do not intersect in their interiors

i. s_n is the number of sequences $i_1 i_2 \cdots i_k$, where $i_j \in \mathbb{P}$ or $i_j = -1$ (and k can be arbitrary), such that $n = \#\{j : i_j = -1\}$, $i_1 + i_2 + \cdots + i_j \geq 0$ for all j, and $i_1 + i_2 + \cdots + i_k = 0$.

j. r_n is the number of lattice paths from $(0, 0)$ to (n, n), with steps $(1, 0)$, $(0, 1)$, and $(1, 1)$, that never rise above the line $y = x$.

k. r_{n-1} is the number of $n \times n$ permutation matrices P with the following property: We can eventually reach the all 1's matrix by starting with P and continually replacing a 0 by a 1 if that 0 has at least two adjacent 1's, where an entry a_{ij} is defined to be adjacent to $a_{i\pm1,j}$ and $a_{i,j\pm1}$.

l. Let $u = u_1 \cdots u_k \in \mathfrak{S}_k$. We say that a permutation $w = w_1 \cdots w_n \in \mathfrak{S}_n$ is *u-avoiding* if no subsequence $w_{a_1}, \ldots, w_{a_k}$ (with $a_1 < \cdots < a_k$) is in the same relative order as u, i.e., $u_i < u_j$ if and only if $w_{a_i} < w_{a_j}$. Let $\mathfrak{S}_n(u, v)$ denote the set of permutations $w \in \mathfrak{S}_n$ avoiding both the permutations $u, v \in \mathfrak{S}_4$. There is a group G of order 16 that acts on the set of pairs (u, v) of unequal elements of $\mathfrak{S}_4$ such that if (u, v) and (u', v') are in the same G-orbit (in which case we say that they are *equivalent*), then there is a simple bijection between $\mathfrak{S}_n(u, v)$ and $\mathfrak{S}_n(u', v')$ (for all n). Namely, identifying a permutation with the corresponding permutation matrix, the orbit of (u, v) is obtained by possibly interchanging u and v, and then doing a simultaneous dihedral symmetry of the square matrices u and v. There are then ten inequivalent pairs $(u, v) \in \mathfrak{S}_4 \times \mathfrak{S}_4$ for which $\#\mathfrak{S}_n(u, v) = r_{n-1}$, namely,

(1234, 1243), (1243, 1324), (1243, 1342), (1243, 2143), (1324, 1342), (1342, 1423), (1342, 1432), (1342, 2341), (1342, 3142), and (2413, 3142).

m. r_{n-1} is the number of permutations $w = w_1 w_2 \cdots w_n$ of $[n]$ with the following property: It is possible to insert the numbers $w_1, \ldots, w_n$ in order into a string, and to remove the numbers from the string in the order $1, 2, \ldots, n$. Each insertion must be at the beginning or end of the string. At any time we may remove the first (leftmost) element of the string. (*Example:* $w = 2413$. Insert 2, insert 4 at the right, insert 1 at the left, remove 1, remove 2, insert 3 at the left, remove 3, remove 4.)

n. r_n is the number of sequences of length $2n$ from the alphabet A, B, C such that: (i) for every $1 \le i < 2n$, the number of A's and B's among the first i terms is not less than the number of C's, (ii) the total number of A's and B's is n (and hence the also the total number of C's), and (iii) no two consecutive terms are of the form *CB*.

o. r_{n-1} is the number of noncrossing partitions (as defined in Exercise 3.68) of some set $[k]$ into n blocks, such that no block contains two consecutive integers.

p. s_n is the number of graphs G (without loops and multiple edges) on the vertex set $[n+2]$ with the following two properties: (α) All of the edges $\{1, n+2\}$ and $\{i, i+1\}$ are edges of G, and (β) G is *noncrossing*, i.e., there are not both edges $\{a, c\}$ and $\{b, d\}$ with $a < b < c < d$. Note that an arbitrary noncrossing graph on $[n+2]$ can be obtained from those satisfying (α)–(β) by deleting any subset of the required edges in (α). Hence the total number of noncrossing graphs on $[n+2]$ is $2^{n+2} s_n$.

q. r_{n-1} is the number of reflexive and symmetric relations R on the set $[n]$ such that if iRj with $i < j$, then we never have uRv for $i \le u < j < v$.

r. r_{n-1} is the number of reflexive and symmetric relations R on the set $[n]$ such that if iRj with $i < j$, then we never have uRv for $i < u \le j < v$.

s. r_{n-1} is the number of ways to cover with disjoint dominos (or dimers) the set of squares consisting of $2i$ squares in the i-th row for $1 \le i \le n-1$, and with $2(n-1)$ squares in the n-th row, such that the row centers lie on a vertical line. See Figure 6-9 for the case $n = 4$.

6.40. [3–] Let a_n be the number of permutations $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$ such that we never have $w_{i+1} = w_i \pm 1$, e.g., $a_4 = 2$, corresponding to 2413 and 3142.

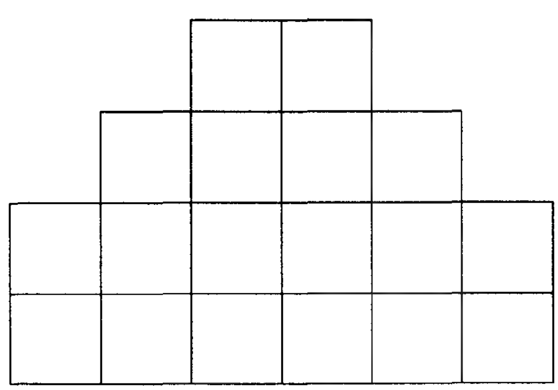


Figure 6-9. A board with $r_3 = 22$ domino tilings.

Equivalently, a_n is the number of ways to place n nonattacking kings on an $n \times n$ chessboard with one king in every row and column. Let

$$\begin{aligned} A(x) &= \sum_{n\geq 0} a_n x^n \\ &= 1 + x + 2x^4 + 14x^5 + 90x^6 + 646x^7 + 5242x^8 + \cdots. \end{aligned}$$

Show that $A(xR(x)) = \sum_{n\geq 0} n!x^n := E(x)$, where

$$R(x) = \sum_{n\geq 0} r_n x^n = \frac{1}{2x}(1 - x - \sqrt{1 - 6x + x^2}),$$

the generating function for Schröder numbers. Deduce that

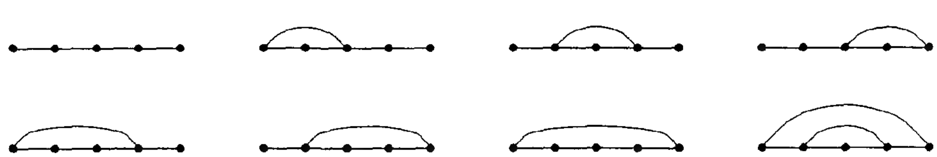
$$A(x) = E\left(\frac{x(1-x)}{1+x}\right).$$

6.41. [3] A permutation $w \in \mathfrak{S}_n$ is called *2-stack sortable* if $S^2(w) = w$, where S is the operator of Exercise 6.19(ii). Show that the number $S_2(n)$ of 2-stack sortable permutations in $\mathfrak{S}_n$ is given by

$$S_2(n) = \frac{2(3n)!}{(n+1)!\,(2n+1)!}.$$

6.42. [2] A king moves on the vertices of the infinite chessboard $\mathbb{Z} \times \mathbb{Z}$ by stepping from (i, j) to any of the eight surrounding vertices. Let $f(n)$ be the number of ways in which a king can walk from $(0, 0)$ to $(n, 0)$ in n steps. Find $F(x) = \sum_{n\geq 0} f(n)x^n$, and find a linear recurrence with polynomial coefficients satisfied by $f(n)$.

6.43. **a.** [2+] A *secondary structure* is a graph (without loops or multiple edges) on the vertex set $[n]$ such that (a) $\{i, i+1\}$ is an edge for all $1 \leq i \leq n-1$, (b) for all i, there is at most one j such that $\{i, j\}$ is an edge and $|j - i| \neq 1$, and (c) if $\{i, j\}$ and $\{k, l\}$ are edges with $i < k < j$, then $i \leq l \leq j$. (Equivalently, a secondary structure may be regarded as a 3412-avoiding involution (as in Exercise 6.19(kk)) such that no orbit consists of two consecutive integers.) Let $s(n)$ be the number of secondary structures with n vertices. For instance, $s(5) = 8$, given by



Let $S(x) = \sum_{n\geq 0} s(n)x^n = 1 + x + x^2 + 2x^3 + 4x^4 + 8x^5 + 17x^6 + 37x^7 + 82x^8 + 185x^9 + 423x^{10} + \cdots$. Show that

$$S(x) = \frac{x^2 - x + 1 - \sqrt{1 - 2x - x^2 - 2x^3 + x^4}}{2x^2}.$$

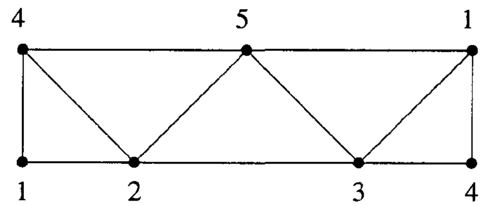


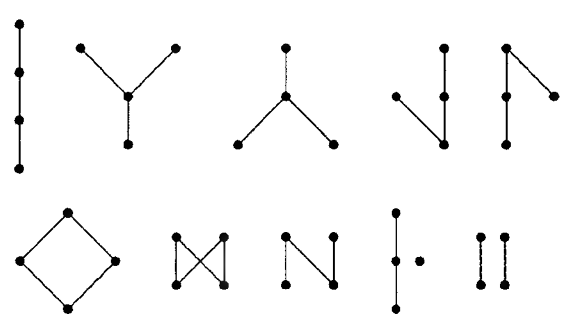
Figure 6-10. A Catalan triangulation of the Möbius band.

b. [3−] Show that $s(n)$ is the number of walks in n steps from $(0, 0)$ to the x-axis, with steps $(1, 0)$, $(0, 1)$, and $(0, -1)$, never passing below the x-axis, such that $(0, 1)$ is never followed directly by $(0, -1)$.

6.44. [3−] Define a *Catalan triangulation* of the Möbius band to be an abstract simplicial complex triangulating the Möbius band that uses no interior vertices, and has vertices labeled $1, 2, \ldots, n$ in order as one traverses the boundary. (If we replace the Möbius band by a disk, then we get the triangulations of Corollary 6.2.3(vi) or Exercise 6.19(a).) Figure 6-10 shows the smallest such triangulation, with five vertices (where we identify the vertical edges of the rectangle in opposite directions). Let $\mathrm{MB}(n)$ be the number of Catalan triangulations of the Möbius band with n vertices. Show that

$$\sum_{n\geq 0} \mathrm{MB}(n)x^n = \frac{x^2[(2-5x-4x^2)+(-2+x+2x^2)\sqrt{1-4x}]}{(1-4x)[1-4x+2x^2+(1-2x)\sqrt{1-4x}]}$$
$$= x^5+14x^6+113x^7+720x^8+4033x^9+20864x^{10}+\cdots.$$

6.45. [3−] Let $f(n)$ be the number of nonisomorphic n-element posets with no 3-element antichain. For instance, $f(4) = 10$, corresponding to



Let $F(x) = \sum_{n\geq 0} f(n)x^n = 1 + x + 2x^2 + 4x^3 + 10x^4 + 26x^5 + 75x^6 + 225x^7 + 711x^8 + 2311x^9 + 7725x^{10} + \cdots$. Show that

$$F(x) = \frac{4}{2-2x+\sqrt{1-4x}+\sqrt{1-4x^2}}.$$

6.46. a. [3+] Let $f(n)$ denote the number of subsets S of $\mathbb{N} \times \mathbb{N}$ of cardinality n with the following property: If $p \in S$ then there is a lattice path from $(0, 0)$

to p with steps $(0, 1)$ and $(1, 0)$, all of whose vertices lie in S. Show that

$$\sum_{n\geq 1} f(n)x^n = \frac{1}{2}\left(\sqrt{\frac{1+x}{1-3x}} - 1\right)$$
$$= x + 2x^2 + 5x^3 + 13x^4 + 35x^5 + 96x^6 + 267x^7 + 750x^8 + 2123x^9 + 6046x^{10} + \cdots.$$

b. [3+] Show that the number of such subsets contained in the first octant $0 \leq x \leq y$ is the Motzkin number M_{n-1} (defined in Exercise 6.37).

6.47. **a.** [3] Let P_n be the Bruhat order on the symmetric group $\mathfrak{S}_n$ as defined in Exercise 3.75(a). Show that the following two conditions on a permutation $w \in \mathfrak{S}_n$ are equivalent:

(i) The interval $[\hat{0}, w]$ of P_n is rank-symmetric, i.e., if ρ is the rank function of P_n (so $\rho(w)$ is the number of inversions of w), then

$$\#\{u \in [\hat{0}, w] : \rho(u) = i\} = \#\{u \in [\hat{0}, w] : \rho(w) - \rho(u) = i\},$$

for all $0 \leq i \leq \rho(w)$.

(ii) The permutation $w = w_1 w_2 \cdots w_n$ is 4231 and 3412-avoiding, i.e., there do not exist $a < b < c < d$ such that $w_d < w_b < w_c < w_a$ or $w_c < w_d < w_a < w_b$.

b. [3−] Call a permutation $w \in \mathfrak{S}_n$ *smooth* if it satisfies (i) (or (ii)) above. Let $f(n)$ be the number of smooth $w \in \mathfrak{S}_n$. Show that

$$\sum_{n\geq 0} f(n)x^n = \frac{1}{1 - x - \frac{x^2}{1-x}\left(\frac{2x}{1+x-(1-x)C(x)} - 1\right)}$$
$$= 1 + x + 2x^2 + 6x^3 + 22x^4 + 88x^5 + 366x^6 + 1552x^7 + 6652x^8 + 28696x^9 + \cdots,$$

where $C(x) = (1 - \sqrt{1-4x})/2x$ is the generating function for the Catalan numbers.

6.48. [3] Let $f(n)$ be the number of 1342-avoiding permutations $w = w_1 w_2 \cdots w_n$ in $\mathfrak{S}_n$, i.e., there do not exist $a < b < c < d$ such that $w_a < w_d < w_b < w_c$. Show that

$$\sum_{n\geq 0} f(n)x^n = \frac{32x}{1 + 20x - 8x^2 - (1-8x)^{3/2}}$$
$$= 1 + x + 2x^2 + 6x^3 + 23x^4 + 103x^5 + 512x^6 + 2740x^7 + 15485x^8 + \cdots.$$

6.49. **a.** [3−] Let B_n denote the board consisting of the following number of squares in each row (read top to bottom), with the centers of the rows lying on a

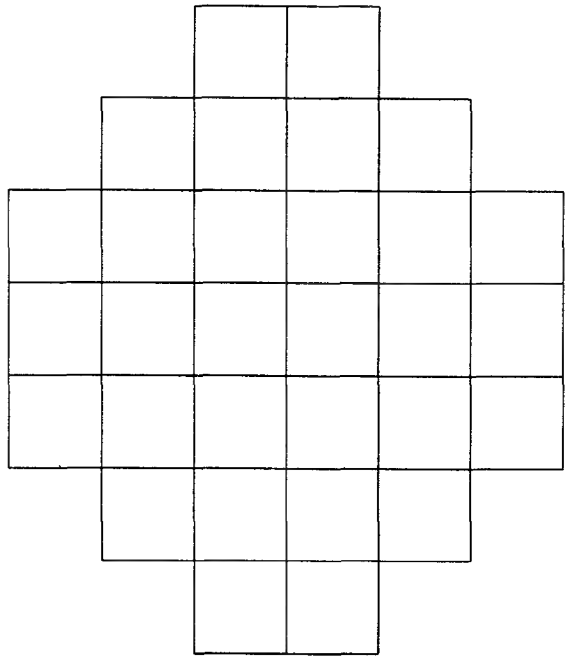


Figure 6-11. A board with $D(3, 3) = 63$ domino tilings.

vertical line: $2, 4, 6, \ldots, 2(n-1), 2n$ (three times), $2(n-1), \ldots, 6, 4, 2$. Figure 6-11 shows the board B_3. Let $f(n)$ be the number of ways to cover B_n with disjoint dominos (or dimers). (A domino consists of two squares with an edge in common.) Show that $f(n)$ is equal to the central Delannoy number $D(n, n)$ (as defined in Section 6.3).

b. [3−] What happens when there are only two rows of length $2n$?

6.50. [3] Let B denote the "chessboard" $\mathbb{N} \times \mathbb{N}$. A *position* consists of a finite subset S of B, whose elements we regard as *pebbles*. A *move* consists of replacing some pebble, say at cell (i, j), with two pebbles at cells $(i+1, j)$ and $(i, j+1)$, provided that each of these cells is not already occupied. A position S is *reachable* if there is some sequence of moves, beginning with a single pebble at the cell $(0, 0)$, that terminates in the position S. A subset T of B is *unavoidable* if every reachable set intersects T. A subset T of B is *minimally unavoidable* if T is unavoidable, but no proper subset of T is unavoidable. Let $u(n)$ be the number of n-element minimally unavoidable subsets of B. Show that

$$\sum_{n\geq 0} u(n)x^n = x^3\frac{(1-3x+x^2)\sqrt{1-4x}-1+5x-x^2-6x^3}{1-7x+14x^2-9x^3}$$

$$= 4x^5 + 22x^6 + 98x^7 + 412x^8 + 1700x^9 + 6974x^{10} + 28576x^{11} + \cdots .$$

6.51. [3+] Let E_n denote the expected number of real eigenvalues of a random $n \times n$ real matrix whose entries are independent random variables from a standard

(mean zero, variance one) normal distribution. Show that

$$\sum_{n\geq 0} E_n x^n = \frac{x(1-x+x\sqrt{2-2x})}{(1-x)^2(1+x)}.$$

6.52. [2] Let b_n be the number of ways of parenthesizing a string of n letters, subject to a *commutative* (but nonassociative) binary operation. Thus for instance $b_5 = 3$, corresponding to the parenthesizations

$$x^2\cdot x^3 \qquad x\cdot(x\cdot x^3) \qquad x(x^2\cdot x^2).$$

(Note that x^3 is unambiguous, since $x\cdot x^2 = x^2\cdot x$.) Let

$$\begin{aligned} B(x) &= \sum_{n\geq 1} b_n x^n \\ &= x + x^2 + x^3 + 2x^4 + 3x^5 + 6x^6 + 11x^7 \\ &\quad + 23x^8 + 46x^9 + 98x^{10} + \cdots. \end{aligned}$$

Show that $B(x)$ satisfies the functional equation

$$B(x) = x + \tfrac{1}{2}B(x)^2 + \tfrac{1}{2}B(x^2). \qquad (6.61)$$

6.53. [2−] Let $n \in \mathbb{P}$, and define

$$f(n) = 1! + 2! + \cdots + n!.$$

Find polynomials $P(x)$ and $Q(x)$ such that

$$f(n+2) = P(n)f(n+1) + Q(n)f(n)$$

for all $n \geq 1$.

6.54. **a.** [2] Fix $r \in \mathbb{P}$, and define

$$S_n^{(r)} = \sum_{k=0}^{n} \binom{n}{k}^r.$$

Show that the function $S_n^{(r)}$ is P-recursive (as a function of n). More generally, if $d \in \mathbb{P}$ is also fixed, let

$$S_n^{(r,d)} = \sum_{\substack{a_1+\cdots+a_d=n \\ a_i\in\mathbb{N}}} \binom{n}{a_1,\ldots,a_d}^r.$$

Then $S_n^{(r,d)}$ is P-recursive.

b. [3−] Show that

$$\begin{aligned} S_{n+1}^{(1)} - 2S_n^{(1)} &= 0 \\ (n+1)S_{n+1}^{(2)} - (4n+2)S_n^{(2)} &= 0 \\ (n+1)^2 S_{n+1}^{(3)} - (7n^2+7n+2)S_n^{(3)} - 8n^2 S_{n-1}^{(3)} &= 0 \\ (n+1)^3 S_{n+1}^{(4)} - 2(6n^3+9n^2+5n+1)S_n^{(4)} - (4n+1)(4n)(4n-1)S_{n-1}^{(4)} &= 0. \end{aligned}$$

c. [3] Show that in fact $S_n^{(r)}$ satisfies a homogeneous linear recurrence of order $\lfloor \frac{r+1}{2} \rfloor$ with polynomial coefficients.

6.55. a. [3] A *Baxter permutation* (originally called a *reduced* Baxter permutation) is a permutation $w \in \mathfrak{S}_n$ satisfying: if $w(r) = i$ and $w(s) = i + 1$, then there is a k_i between r and s (i.e., $r \le k_i \le s$ or $s \le k_i \le r$) such that $w(t) \le i$ if t is between r and k_i, while $w(t) \ge i+1$ if $k_i+1 \le t \le s$ or $s \le t \le k_i - 1$. For instance, all permutations $w \in \mathfrak{S}_4$ are Baxter permutations except 2413 and 3142. Let $B(n)$ denote the number of Baxter permutations in $\mathfrak{S}_n$. Show that

$$B(n) = \binom{n+1}{1}^{-1}\binom{n+1}{2}^{-1}\sum_{k=1}^{n}\binom{n+1}{k-1}\binom{n+1}{k}\binom{n+1}{k+1}. \tag{6.62}$$

b. [2+] Deduce that $B(n)$ is P-recursive.

c. [3−] Find a (nonzero) homogeneous linear recurrence with polynomial coefficients satisfied by $B(n)$.

6.56. a. [3] An *increasing subsequence* of length j of a permutation $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$ is a sequence $w_{i_1} w_{i_2} \cdots w_{i_j}$ such that $i_1 < i_2 < \cdots < i_j$ and $w_{i_1} < w_{i_2} < \cdots < w_{i_j}$. Fix $k \in \mathbb{P}$. Let $A_k(n)$ denote the number of permutations $w \in \mathfrak{S}_n$ that have no increasing subsequence of length k. Show that the function A_k is P-recursive.

b. [5] Show that the numbers $A_{k+1}(n)$ satisfy a recurrence of the form

$$\sum_{i=0}^{\lfloor k/2 \rfloor} p_i(n) A_{k+1}(n-i) = 0, \tag{6.63}$$

where $p_i(n)$ is a polynomial of degree at most k with the following additional properties:

(i)

$$p_0(n) = \prod_{i=1}^{\lfloor k/2 \rfloor} [n + i(k-i)]^2.$$

(ii) For $2 \le i \le \lfloor k/2 \rfloor + 1$ we have

$$p_i(n) = q_i(n) \prod_{j=1}^{i-1} (n-j)^2, \tag{6.64}$$

where $q_i(n)$ is a polynomial of degree at most $k - 2i$.

(iii) The polynomials $q_i(n)$ of (6.64) are such that the recurrence (6.63) is true with the unique initial condition $A_{k+1}(0) = 1$.

(iv) If $k = 2m + 1$ then the leading coefficient of $q_i(n)$ is the coefficient of z^i in the polynomial

$$\prod_{j=0}^{m} [1 - (2j+1)^2 z].$$

c. [5] Fix a permutation $u \in \mathfrak{S}_k$. Let $A_u(n)$ denote the number of u-avoiding permutations $w \in \mathfrak{S}_n$, as defined in Exercise 6.39(l). Is A_u P-recursive?

6.57. [3−] Let $f(n)$ for $n \in \mathbb{N}$ satisfy a homogeneous linear recurrence relation of order d with constant coefficients over $\mathbb{C}$, i.e.,

$$f(n+d) + \alpha_1 f(n+d-1) + \cdots + \alpha_d f(n) = 0, \qquad (6.65)$$

and suppose that f satisfies no such recurrence of smaller order. What is the smallest order of a (nonzero) homogeneous linear recurrence relation with *polynomial* coefficients satisfied by f? (The answer will depend on the recurrence (6.65), not just on d.) (*Example:* If $f(n) = n$, then f satisfies $f(n+2) - 2f(n+1) + f(n) = 0$ (constant coefficients) and $nf(n+1) - (n+1)f(n) = 0$ (polynomial coefficients).)

6.58. [2+] Consider the homogeneous linear equation (6.34), where $P_e(n) \neq 0$ and the ground field K has characteristic 0. Let $\mathcal{V}$ be the K-vector space of all solutions $f : \mathbb{N} \to K$ to (6.34). Show that

$$e \leq \dim \mathcal{V} \leq e + m, \qquad (6.66)$$

where m is the number of distinct zeros of $P_e(n)$ that are nonnegative integers. Show that for fixed e and $P_e(n)$, any value of $\dim \mathcal{V}$ in the range given by (6.66) can occur.

6.59. [3−] Show by direct formal arguments that the series $\sec x$ and $\sqrt{\log(1+x^2)}$ are not D-finite. Hence the reciprocal and square root of a D-finite series need not be D-finite.

6.60. [3+] Let $y \in \mathbb{C}[[x]]$ be D-finite with $y(0) \neq 0$. Show that $1/y$ is D-finite if and only if y'/y is algebraic.

6.61. [3+] Let $F(x_1, \ldots, x_m) = \sum_{\alpha \in \mathbb{N}^m} f(\alpha) x^\alpha$ be a rational power series over the field K. Show that

$$\operatorname{diag}(F) := \sum_{n \geq 0} f(n, \ldots, n) t^n$$

is D-finite.

6.62. [3] Let $y \in \mathbb{C}_{\text{alg}}[[x]]$. Thus by Theorem 6.4.6, y satisfies a homogeneous linear differential equation with polynomial coefficients. Show that the least order of such an equation is equal to the dimension of the complex vector space V spanned by y and all its conjugates. (*Example:* Suppose $y^d = R(x) \in \mathbb{C}(x)$, where $y^d - R(x)$ is irreducible over $\mathbb{C}(x)$. The conjugates of y are given by ζy, where $\zeta^d = 1$. Hence y and all its conjugates span a one-dimensional vector space, and equation (6.37) gives the differential equation of order one satisfied by y.)

6.63. **a.** [5] Suppose that $y = \sum_{n \geq 0} a_n x^n \in \mathbb{C}[[x]]$ is D-finite. Define $\chi : \mathbb{C} \to \mathbb{Z}$ by

$$\chi(a) = \begin{cases} 1, & a \neq 0 \\ 0, & a = 0. \end{cases}$$

Is $\sum_{n \geq 0} \chi(a_n) x^n$ rational? This question is open even if y is just assumed to be algebraic; for the case when y is rational see Exercise 4.3.

b. [3] Show that (a) is false if y is assumed only to satisfy an *algebraic differential equation* (ADE), i.e., there is a nonzero polynomial $f(x_1, x_2, \dots, x_{d+2}) \in \mathbb{C}[x_1, x_2, \dots, x_{d+2}]$ such that $f(x, y, y', y'', \dots, y^{(d)}) = 0$.

c. [5] Suppose that y satisfies an ADE and $y \notin \mathbb{C}[x]$. Can y be more than quadratically lacunary? In other words, if $y = \sum b_i x^{n_i}$, can one have $\lim_{i\to\infty} i^2/n_i = 0$?

6.64. a. [2] Let x and y be noncommuting indeterminates. Show that the following identity is valid, in the sense that the formal series represented by both sides are the same.

$$(1+x)(1-yx)^{-1}(1+y) = (1+y)(1-xy)^{-1}(1+x). \quad (6.67)$$

b. [3−] Is equation (6.67) valid in any associative algebra with identity for which both sides are defined?

c. [3] Even more generally, let R be the ring obtained from the noncommutative polynomial ring $K\langle x_1, \dots, x_n\rangle$ by successively joining inverses of all elements whose series expansion has constant term 1. Let ω be the homomorphism from R to the ring $K\langle\langle x_1, \dots, x_n\rangle\rangle$ that replaces a rational "function" by its series expansion. Is ω one-to-one?

6.65. a. [2+] Verify the statement preceding Definition 6.5.3 and in Example 6.6.2 that the series $\sum_{n\geq 1} x^n y^n$ is not rational.

b. [3−] More generally, suppose that $S \in K_{\text{rat}}\langle\langle X\rangle\rangle$. Let supp$(S)$ be defined as in Definition 6.6.4. Show that there is a $p \in \mathbb{P}$ such that every $z \in \text{supp}(S)$ of length at least p can be decomposed as $z = uvw$ where $v \neq 1$ and $\{uv^n w : n \geq 0\} \cap \text{supp}(S)$ is infinite.

c. [3] Let $L \subseteq X^*$ be a language. Show that L is rational (as defined in Definition 6.6.4) if and only if there is a $p \in \mathbb{P}$ satisfying the following condition: Every word $x \in X^*$ of length p can be decomposed as $x = uvw$, where $v \neq 1$ and, for all $y \in X^*$ and all $n \in \mathbb{N}$, we have that $xy \in L$ if and only if $uv^n wy \in L$.

6.66. [2+] Let $S, T \in K\langle\langle X\rangle\rangle$ be rational series. Show that the Hadamard product $S * T$ (as defined in Section 6.6) is also rational.

6.67. [2] Let $b \geq 2$. If $n = i_k b^k + \cdots + i_1 b + i_0$ is the base b expansion of n (with $i_k > 0$), then associate with n the word $w(n) = x_{i_k} \cdots x_{i_1} x_{i_0}$ in the alphabet $X = \{x_0, \dots, x_{b-1}\}$. Let $S = \sum_{n\geq 0} n w(n)$, the generating function for base b expansions of positive integers. For instance, if $b = 2$ then

$$S = x_1 + 2x_1x_0 + 3x_1^2 + 4x_1x_0^2 + 5x_1x_0x_1 + 6x_1^2x_0 + 7x_1^3 + \cdots.$$

Show that S is rational.

6.68. [3+] Let (W, S) be a Coxeter group, i.e., W is generated by $S = \{s_1, \dots, s_n\}$ subject to relations $s_i^2 = 1$, $(s_i s_j)^{m_{ij}} = 1$ for all $i < j$, where $m_{ij} \geq 2$ (possibly $m_{ij} = \infty$, meaning there is no relation $(s_i s_j)^{m_{ij}} = 1$). A word $x_{i_1} x_{i_2} \cdots x_{i_p}$ in the alphabet $X = \{x_1, \dots, x_n\}$ is *reduced* if there is no relation $s_{i_1} s_{i_2} \cdots s_{i_p} = s_{j_1} s_{j_2} \cdots s_{j_q}$ in W with $q < p$. Show that the set of reduced words is a rational language, as defined in Definition 6.6.4.

6.69. [3−] Let A be a finite alphabet. Given $u, v \in A^*$, define $u \le v$ if u is a subword of v. Write $u \otimes v$ for elements of the direct product $A^* \times A^*$. Show that the series $\sum_{u \le v} u \otimes v$, and hence the language $\{u \otimes v \in A^* \times A^* : u \le v\}$, is rational.

6.70. [1] Find explicitly the solution (R_1, R_2) to the proper algebraic system

$$\begin{aligned} z_1 &= y^2 z_1 - z_2^2 \\ z_2 &= z_2 x z_1 + z_1 z_2. \end{aligned}$$

6.71. [2] Let X be a finite alphabet. Show that the set $K_{\text{alg}}\langle\langle X\rangle\rangle$ of all algebraic series in X forms a subalgebra of $K\langle\langle X\rangle\rangle$. Moreover, if $u \in K_{\text{alg}}\langle\langle X\rangle\rangle$ and u^{-1} exists, then $u^{-1} \in K_{\text{alg}}\langle\langle X\rangle\rangle$.

6.72. **a.** [3−] Let $L \in K_{\text{alg}}\langle\langle X\rangle\rangle$ be an algebraic language, as defined in Definition 6.6.4. Show that there is a $p \in \mathbb{P}$ such that every $z \in L$ of length at least p can be decomposed as $z = uvwxy$ where $vx \ne 1$ and $\{uv^n w x^n y : n \ge 0\} \cap L$ is infinite.

b. [1+] Deduce from (a) that $\sum_{n \ge 0} x^n y^n z^n$ is not algebraic.

6.73. **a.** [2+] Let $S = \sum w$, summed over all words $w \in \{x, y\}^*$ with the same number of x's as y's. Thus

$$\phi(S) = \sum_{n \ge 1} \binom{2n}{n} x^n y^n = \frac{1}{\sqrt{1 - 4xy}},$$

where ϕ is the "abelianization" operator. Show that S is algebraic.

b. [5−] What is the least number of equations in a proper algebraic system for which $S - 1$ is a component?

6.74. [3−] Let $x_1, \ldots, x_k$ be noncommuting indeterminates. Let $f(n)$ be the constant term of the noncommutative Laurent polynomial $(x_1 + x_1^{-1} + \cdots + x_k + x_k^{-1})^n$. Show that

$$\sum_{n \ge 0} f(n) t^n = \frac{2k - 1}{k - 1 + k\sqrt{1 - 4(2k - 1)t^2}},$$

as stated in Example 6.7.3.

Solutions to Exercises

6.1. One way to proceed is as follows. Suppose that e^x is algebraic of degree d, and let

$$P_d(x) e^{dx} + P_{d-1}(x) e^{(d-1)x} + \cdots + P_0(x) = 0, \tag{6.68}$$

where $P_i(x) \in \mathbb{C}[x]$, and $\deg P_d(x)$ is minimal. Differentiate (6.68) with respect to x and subtract equation (6.68) multiplied by d. We get the equation

$$(P_0' - dP_0) + [P_1' - (d-1)P_1]e^x + \cdots + (P_{d-1}' - P_{d-1})e^{(d-1)x} + P_d' e^{dx} = 0,$$

which either has degree less than d or else contradicts the minimality of $\deg P_d$. For a less straightforward but also less *ad hoc* solution, see Exercise 6.2.

6.2. **a.** This result is known as *Eisenstein's theorem.* For a proof and additional references, see G. Pólya and G. Szegö, *Problems and Theorems in Analysis*, vol. II, Springer-Verlag, New York/Berlin/Heidelberg, 1976 (Part 8, Ch. 3, §§2–3).

b. Easy from (a).

6.3. It follows from a theorem of R. Jungen [**40**] that if $y = \sum a_n x^n \in \mathbb{C}[[x]]$ is algebraic and $a_n \sim cn^r\alpha^n$ for some constants $0 \neq c \in \mathbb{C}$, $0 > r \in \mathbb{R}$, and $0 \neq \alpha \in \mathbb{C}$, then $r = s + \frac{1}{2}$ for some $s \in \mathbb{Z}$. Since by Stirling's formula $\binom{3n}{n,n,n} \sim cn^{-1}\alpha^n$ and $\binom{2n}{n}^2 \sim dn^{-1}\beta^n$, the proof follows.

For another proof that y_1 and y_2 aren't algebraic, we use the Gaussian hypergeometric series

$$F(a,b,c;x) = \sum_{n\geq 0} \frac{z^n}{n!} \prod_{i=0}^{n-1} \frac{(a+i)(b+i)}{c+i}.$$

It is easy to see that $y_1 = F\left(\frac{1}{3}, \frac{2}{3}, 1; 27x\right)$ and $y_2 = F\left(\frac{1}{2}, \frac{1}{2}, 1; 16x\right)$. It was determined by H. A. Schwarz, *J. Reine Angew. Math.* **75** (1873), 292–335; *Ges. Math. Abh. Bd. II*, Chelsea, New York, 1972, pp. 211–259, for exactly what parameters a, b, c is $F(a,b,c;x)$ algebraic. From this result it follows that y_1 and y_2 aren't algebraic. Alternatively, we could use a result of T. Schneider, *Einführung in die Tranzendenten Zahlen*, Springer, Berlin, 1957, which implies that $y_1(x)$ and $y_2(x)$ take nonalgebraic values (over $\mathbb{Q}$) when $x \neq 0$ is algebraic. Since it is easy to see that algebraic power series with rational coefficients satisfy a polynomial equation with rational coefficients, it follows that y_1 and y_2 aren't algebraic. For yet another proof that y_1 isn't algebraic (due to J. Rickard), see G. Almkvist, W. Dicks, and E. Formanek, *J. Algebra* **93** (1985), 189–214 (p. 209).

Finally, in the paper C. F. Woodcock and H. Sharif, *J. Algebra* **121** (1989), 364–369, appears a clever proof that $z := \sum_{n\geq 0} \binom{2n}{n}^t x^n$ isn't algebraic for any $t \in \mathbb{N}$, $t > 1$. (The asymptotic method of the first proof above works only for even t.) They use the fact that z is algebraic in characteristic $p > 0$ (for which they give a direct proof, though one could also use Exercise 6.11) and show that the degree of z over $\mathbb{F}_p(x)$ is unbounded as $p \to \infty$. From this it follows easily that z is not algebraic. Woodcock and Sharif show that some similar series are not algebraic by the same technique.

In general it seems difficult to determine whether some "naturally occurring" power series y is algebraic. If y is D-finite and one is given the linear differential equation $\mathcal{D}$ of least degree with polynomial coefficients that y satisfies, then a decision procedure for deciding whether y is algebraic (in which case all solutions to $\mathcal{D}$ are algebraic) was given by M. F. Singer, in *Proceedings of the 1979 Queens University Conference on Number Theory*, Queens Papers in Pure and Applied Mathematics **54**, pp. 378–420. (The problem had been almost completely solved by Boulanger and Painlevé in the nineteenth century, as discussed in the previous reference.) Singer has improved and generalized his result in several papers; for a brief discussion and references see his paper in *Computer Algebra and Differential Equations* (E. Tournier, ed.), Academic Press, New York, 1990, pp. 3–57.

* **6.4.** Let K be the algebraic closure of $\mathbb{F}_p$, the finite field of order p. Then the equation $y^p - y - x^{-1}$ has no solution that belongs to $K^{\text{fra}}((x))$, as may be seen by

considering the smallest integer $N > 0$ for which $y \in K((x))[x^{1/N}]$. This example is due to Chevalley [**12**, § IV.6]. Subsequently S. Abhyankar, *Proc. Amer. Math. Soc.* **7** (1956), 903–905, gave the factorization

$$y^p - y - x^{-1} = \prod_{i=0}^{p-1}\left(y - i - \sum_{j\geq 1} x^{-1/p^j}\right).$$

Using this example Men-Fon Huang, Ph.D. thesis, Purdue University, 1968, constructed a "generalized Puiseux field" $A(p)$ containing an algebraic closure of $K((x))$. Further properties of $A(p)$ were developed by S. Vaidya, *Illinois J. Math.* **41** (1997), 129–141. Another example of an equation with no solution in $K^{\mathrm{fra}}((x))$ is mentioned by Cohn [**17**, p. 198] and attributed to M. Ojanguren.

6.5. Let $\eta_1, \dots, \eta_d$ be the roots of P, so by Proposition 6.1.8 we have $\eta_1, \dots, \eta_d \in K^{\mathrm{fra}}[[x]]$. Suppose $\eta_j = \sum_{n\geq 0} a_n x^{n/N}$, where $\deg_{K((x))}\eta_j = N$ (so $N = c_s$ for some s). Let ζ be a primitive N-th root of unity, and set $\eta_j^{(k)} = \sum_{n\geq 0} a_n \zeta^{kn} x^{n/N}$ for $0 \leq k \leq N$. Thus disc $P(y)$ is divisible by $\prod_{0\leq a<b\leq N-1}\left(\eta_j^{(a)} - \eta_j^{(b)}\right)^2$. Each $\eta_j^{(a)} - \eta_j^{(b)}$ has zero constant term and hence is divisible by $x^{1/N}$, so the product is divisible by $(x^{1/N})^{2\binom{N}{2}} = x^{N-1}$. The stated result follows easily.

6.6. Follows easily from Exercise 1.37(a).

6.7. Yes. If u satisfies the polynomial equation $P(x, u) = 0$, then $u^{\langle -1\rangle}$ satisfies $P(u^{\langle -1\rangle}, x) = 0$.

6.8. **a.** The discriminant of a polynomial $F(y)$ is 0 if and only if $F(y)$ and $F'(y)$ have a common nonconstant factor. If $F(y) = ay^d + by + c$, then

$$dF(y) - yF'(y) = b(d-1)y + cd.$$

Assume $c \neq 0$. (Otherwise the problem is easy.) Then disc $F(y) = 0$ if and only if $F'(y)$ is divisible by $b(d-1)y + cd$, i.e., $F'(-cd/b(d-1)) = 0$. This leads to the condition

$$d^d a c^{d-1} + (-1)^{d-1}(d-1)^{d-1} b^d = 0,$$

and a simple normalization argument yields equation (6.7).

A somewhat different approach is the following. It is easily seen that for *any* polynomial $G(y) = ay^d + \cdots = a(y - y_1)\cdots(y - y_d)$, we have

$$\operatorname{disc} G = (-1)^{\binom{d}{2}} a^{d-2} G'(y_1)\cdots G'(y_d). \tag{6.69}$$

Letting $G(y) = F(y)$, we see that $F'(y_i) = ady_i^{d-1} + b$, so $y_i F'(y_i) = ady_i^d + by_i = d(-by_i - c) + by_i = b(1-d)y_i - cd$. Hence

$$\begin{aligned}\operatorname{disc} F &= (-1)^{\binom{d}{2}} a^{2d-2} \prod_{i=1}^{d} \frac{b(1-d)y_i - cd}{y_i}\\ &= (-1)^{\binom{d}{2}} a^{2d-2} \frac{b^d(d-1)^d a^{-1} F(cd/b(1-d))}{(-1)^d a^{-1} c},\end{aligned}$$

which by routine manipulation becomes (6.7). This result appears in G. Salmon, *Lectures Introductory to the Modern Higher Algebra*, Dublin, 1885, reprinted by Chelsea, New York, 1969, and E. Netto, *Vorlesungen über Algebra*, Teubner-Verlag, Leipzig, 1896, though it may go back earlier. See also [**28**, Ch. 12, (1.38)].

b. Let

$$P = \frac{x^n}{n!} + \frac{x^{n-1}}{(n-1)!} + \cdots + x + 1 = \frac{1}{n!}(x - x_1)\cdots(x - x_n).$$

Then $P = x^n/n! + P'$, so by (6.69) we have

$$\begin{aligned}\operatorname{disc} P &= (-1)^{\binom{n}{2}} n!^{-(n-2)} \prod_{i=1}^{n}\left(P(x_i) - \frac{x_i^n}{n!}\right)\\ &= (-1)^{\binom{n}{2}} n!^{-(n-2)}(-1)^n n!^{-n}\left(\prod_{i=1}^{n} x_i\right)^n\\ &= (-1)^{\binom{n-1}{2}} n!^{-2n+2}((-1)^n n!)^n\\ &= (-1)^{\binom{n}{2}} n!^{-(n-2)}.\end{aligned}$$

This result follows from a more general result of D. Hilbert, *J. Reine Angew. Math.* **103** (1888), 337–345, though as in (a) it may have been known earlier. See also [**28**, Ch. 12, (1.42)].

6.9. **a.** Let $*_x$ denote Hadamard product with respect to x only, and let diag denote the diagonal (in the variable y) with respect to the variables y_1 and y_2. It is clear that

$$F * G = \operatorname{diag}(F(x, y_1) *_x F(x, y_2)). \tag{6.70}$$

Now $F(x, y_1)$ and $F(x, y_2)$ are both rational series in x over the field $K(y_1, y_2)$, so by Proposition 4.2.5 we have

$$F(x, y_1) *_x F(x, y_2) \in K(y_1, y_2)(x) = K(x)(y_1, y_2).$$

Then by Theorem 6.3.3, $\operatorname{diag}(F(x, y_1) *_x F(x, y_2))$ is algebraic over $K(x)(y) = K(x, y)$, as desired.

An explicit statement of the result of this exercise first appears in C. F. Woodcock and H. Sharif, *J. Algebra* **128** (1990), 517–527 (Thm. 5.1), who give a more complicated proof than ours. For the case when char $K > 0$, see Exercise 6.11.

b. Write $*_i$ for Hadamard product with respect to x_i, and $*_{ij}$ for Hadamard product with respect to x_i and x_j. Since $(1 - x - y)^{-1} = \sum \binom{m+n}{m} x^m y^n$, it is easy to see that

$$\begin{aligned}F_k = \Bigg[&\frac{1}{1 - x_1 - x_2} *_2 \frac{1}{1 - x_2 - x_3} *_3 \frac{1}{1 - x_3 - x_4}\\ &*_4 \cdots *_{k-1} \frac{1}{1 - x_{k-1} - x_k}\Bigg] *_{1k} \frac{1}{1 - x_1 - x_k}.\end{aligned} \tag{6.71}$$

By successive applications of Proposition 4.2.5, the expression within brackets is rational, so by (a) we have that F_k is algebraic.

c. By (6.70) we need to compute

$$\mathrm{diag}\left(\frac{1}{1-x-y_1} *_x \frac{1}{1-x-y_2}\right).$$

Now

$$\frac{1}{1-x-y_i} = \frac{1}{1-y_i}\cdot\frac{1}{1-\frac{x}{1-y_i}} = \frac{1}{1-y_i}\sum_{n\geq 0}(1-y_i)^{-n}x^n.$$

Hence

$$\begin{aligned}\frac{1}{1-x-y_1} * \frac{1}{1-x-y_2} &= \frac{1}{(1-y_1)(1-y_2)}\sum_{n\geq 0}(1-y_1)^{-n}(1-y_2)^{-n}x^n\\ &= \frac{1}{1-x}\cdot\frac{1}{1-\frac{y_1+y_2-y_1y_2}{1-x}}.\end{aligned}$$

Now apply Exercise 6.15 to obtain

$$F_2 = \frac{1}{\sqrt{(1-x+y)^2-4y}} = \frac{1}{\sqrt{(1-x-y)^2-4xy}}.$$

By similar reasoning we obtain

$$F_3 = \frac{1}{\sqrt{(1-x-y-z)^2-4xyz}}.$$

The results of (b) and (c) are due to L. Carlitz, *SIAM Review* **6** (1964), 20–30, essentially by the same methods as here. Carlitz did not appeal to (a) directly to show (b), but rather found an explicit (though based on a recurrence) formula for F_k. In particular, $F_k = P_k^{-1/2}$ where P_k is a polynomial in $x_1, \ldots, x_k$. (This result can also be obtained by a careful analysis of equation (6.71)). For further information and references related to the "cyclic binomial sums" of the type considered here, see V. Strehl, in *Séries Formelles et Combinatoire Algébrique* (P. Leroux and C. Reutenauer, eds.), Publ. du LACIM, vol. 11, Univ. du Québec à Montréal, 1992, pp. 363–377.

6.10. The series y is just the diagonal of

$$\frac{q^{-m}}{1-qP(q)t} = \sum_{n\geq 0} q^{n-m}P(q)^n t^n,$$

so the proof follows from Theorem 6.3.3. (Technically speaking, Theorem 6.3.3 does not apply because we are dealing with a Laurent series and not a power series, but the proof goes through in exactly the same way.) A somewhat more general result was proved by G. Pólya [53].

6.11. a. This result is due to H. Sharif and C. F. Woodcock, *J. London Math. Soc. (2)* **37** (1988), 395–403 (p. 401).

b. Let $G = 1/(1 - x_1 \cdots x_k)$. By (a), $F * G$ is algebraic over $K(x_1, \ldots, x_k)$. Now $F * G = (\operatorname{diag} F)(x_1 \cdots x_k)$ (the diagonal of F in the variable $x_1 \cdots x_k$). From this it is not hard to deduce that $(\operatorname{diag} F)(x)$ is algebraic over $K(x)$. See *ibid.*, Thm. 7.1, for the details.

This result was first proved by P. Deligne, *Invent. Math.* **76** (1983), 129–143, using sophisticated methods. Another elementary proof (in addition to the one of Sharif and Woodcock just sketched) appears in J. Denef and L. Lipshitz, *J. Number Theory* **26** (1987), 46–67 (Prop. 5.1). Earlier it was shown by H. Furstenberg [**25**] that the diagonal of a *rational* power series in several variables over a field of characteristic $p > 0$ is algebraic. An interesting survey of some aspects of power series in characteristic p (as well as some results valid in characteristic 0), including connections with finite automata, appears in L. Lipshitz and A. J. van der Poorten, in *Number Theory* (R. A. Mollin, ed.), de Gruyter, Berlin/New York, 1990, pp. 340–358.

Note that (a) follows from (b), since

$$(F * G)(x_1, \ldots, x_k) = \operatorname{diag}_{y_1 z_1} \cdots \operatorname{diag}_{y_k z_k} F(y_1, \ldots, y_k) G(z_1, \ldots, z_k),$$

where $\operatorname{diag}_{y_i z_i}$ denotes the diagonal with respect to y_i and z_i in the variable x_i. Hence (a) and (b) are essentially equivalent (as noted by J.-P. Allouche, *Sém. Théor. Nombres Bordeaux (2)* **1** (1989), 163–187).

6.12. Let t be a new indeterminate, and write xt for the variables $x_1 t, \ldots, x_m t$, and similarly for yt. Let $*$ denote the usual Hadamard product with respect to the variable t. We may regard the series $F(xt)$ and $G(yt)$ as rational power series in t over the field $K(x, y)$. Moreover,

$$F(x) \heartsuit G(y) = F(xt) * G(yt)|_{t=1}.$$

Now consider Proposition 4.2.5, which is proved for the field $K = \mathbb{C}$. The proof actually shows that over any field K, the Hadamard product $A * B$ of rational power series A and B in one variable is rational over the algebraic closure of K. But since $A * B$ is defined over K, it is easy to see that $A * B$ is in fact rational over K. It therefore follows that

$$F(xt) * G(yt) \in K(x, y)(t),$$

and the proof follows. A similar argument applies to the case when F is rational and G is algebraic, using Proposition 6.1.11.

The product $\heartsuit$ was first defined (using a different notation) by S. Bochner and W. T. Martin, *Ann. Math.* **38** (1938), 293–302. The result of this exercise was stated without proof by M. P. Schützenberger [**65**, p. 885].

6.13. a. The other roots are given by

$$\theta_\zeta = -\frac{1}{k-1} \sum_{n \geq 0} \binom{kn/(k-1)}{n} \zeta^n x^{n/(k-1)}, \tag{6.72}$$

where $\zeta^{k-1} = 1$. By Corollary 6.1.7, it suffices to prove the assertion for

$\zeta = 1$. Let

$$\tau = \theta_1(x^{k-1}) = -\frac{1}{k-1}\sum_{n\geq 0}\binom{kn/(k-1)}{n}x^n.$$

We need to show that

$$k^k x^{k-1}\tau^k = (\tau - 1)[(k-1)\tau + 1]^k. \tag{6.73}$$

Applying Example 6.2.7 to the series $-(k-1)\tau = (1-k)\tau$ (with k replaced by $k/(k-1)$) gives

$$\frac{(1-k)\tau - 1}{1 + \left(\frac{k}{k-1} - 1\right)(1-k)\tau} = x\left(\frac{\frac{k}{k-1}(1-k)\tau}{1 + \left(\frac{k}{k-1} - 1\right)(1-k)\tau}\right)^{\frac{k}{k-1}}.$$

Simple algebraic manipulations show that this equation is equivalent to (6.73). Since the $k-1$ series (6.72) are all conjugate over $\mathbb{C}(x)$ by Corollary 6.1.7 (or Proposition 6.1.6), either $P(y)$ is irreducible or $y = \sum \binom{kn}{n}x^n$ is rational. It is easy to see [why?] that y can't be rational, so irreducibility follows.

b. *Answer:* $(-1)^{\binom{k}{2}}k^{k(k-1)}[k^k x - (k-1)^{k-1}]x^{k-2}$

6.14. The recurrence relation (6.53) and the values $f(i,0) = 2^{-i}$ imply that

$$(x^2 - 2x + y)F(x,y) = -2x + H(y)$$

for some $H(y) \in \mathbb{Q}[[y]]$. The left-hand side vanishes formally when $x = 1 - \sqrt{1-y}$, so hence also the right-hand side. Thus $H(y) = 2(1 - \sqrt{1-y})$, so

$$F(x,y) = \frac{2(1-\sqrt{1-y}) - 2x}{x^2 - 2x + y} = \frac{2}{1 - x + \sqrt{1-y}}.$$

This exercise is due to R. Pemantle. A slightly different result also due to Pemantle is cited in M. Larsen and R. Lyons, *J. Theor. Probability*, to appear.

6.15. Let

$$F(s,t) = [1 - sf(st) - tg(st) - h(st)]^{-1},$$

so

$$\begin{aligned}\operatorname{diag} F &= [u^0]\,\frac{1}{1 - uf(x) - \frac{x}{u}g(x) - h(x)}\\ &= [u^0]\,\frac{-u}{u^2 f - (h-1)u + xg}.\end{aligned}$$

The computation now parallels that of Examples 6.3.4, 6.3.8, and 6.3.9 (which the present exercise generalizes), yielding

$$\operatorname{diag} F = \frac{1}{\sqrt{(1-h)^2 - 4xfg}}.$$

A "naive" proof can also be given by generalizing the solution to Exercise 1.5(b).

6.17. **a.** Because of the conditions on S, no S-path can move from one side of the line $y = x$ to the other without actually landing on the line $y = x$. Let $\mathcal{P}_m$ be the set of all S-paths from $(0, 0)$ to (m, m) that don't just consist of the single step (m, m) and that stay strictly below $y = x$ except at $(0, 0)$ and (m, m). Every S-path from $(0, 0)$ to (n, n) is a unique product (juxtaposition) of S-paths of the types (i) a single step from S of the form (j, j), (ii) elements of $\mathcal{P}_m$ for some m, and (iii) reflections about $y = x$ of elements of $\mathcal{P}_m$ for some m. Every S-path from $(0, 0)$ to (n, n) that doesn't rise above $y = x$ is a unique product of S-paths of types (i) and (ii). Hence if $P(x) = \sum_{m\geq 1}(\#\mathcal{P}_m)x^m$, then

$$G = \frac{1}{1 - K - 2P}, \qquad H = \frac{1}{1 - K - P}.$$

Eliminating P from these two equations gives $H = 2/(1 - K + G^{-1})$, as desired.

b. We have $K(x) = x$, while from Exercise 1.5(b) or Example 6.3.8 we have $G(x) = 1/\sqrt{1 - 6x + x^2}$. Hence

$$H(x) = \frac{2}{1 - x + \sqrt{1 - 6x + x^2}} = \frac{1 - x - \sqrt{1 - 6x + x^2}}{2x},$$

so by the discussion of Schröder's second problem in Section 6.2 we have $h(n) = r_n$.

c. It was observed in the discussion of Schröder's second problem that $r_n/2$ $(= s_n)$ does indeed count dissections of the appropriate kind. However, this observation leads only to a generating function, not combinatorial, proof that $h(n) = r_n$. A combinatorial proof was subsequently given by L. W. Shapiro and R. A. Sulanke, Bijections for the Schröder numbers, preprint.

6.18. See I. M. Gessel, *J. Combinatorial Theory (A)* **28** (1980), 321–337 (Cor. 5.4).

6.19. It would require a treatise in itself to discuss thoroughly all the known interconnections among these problems. We will content ourselves with some brief hints and references that should serve as a means of further exposure to "Catalan disease" or "Catalania" (= Catalan mania). One interesting item omitted from the list because of its complicated description is the number of flexagons of order $n + 1$; see C. O. Oakley and R. J. Wisner, *Amer. Math. Monthly* **64** (1957), 143–154 (esp. p. 152) for more information.

Parts a, b, c, h, i, r are covered by Corollary 6.2.3.

d. This is covered by Example 5.3.12. Alternatively, do a depth-first search, ignoring the root edge, recording 1 when a left edge is first encountered, and recording -1 when a right edge is first encountered. This gives a bijection with (r). Note also that when all endpoints are removed (together with the incident edges), we obtain the trees of (c).

e. This is covered by Example 6.2.8. For a bijection with (r), do a depth-first (preorder) search through the tree. When going "down" an edge (away from the root) record a 1, and when going up an edge record a -1. For further information and references, see D. A. Klarner, *J. Combinatorial Theory* **9** (1970), 401–411.

f. When the root is removed we obtain the trees of (d). See also Klarner, *op. cit.*

g. The bijection between parts (i) and (iv) of Proposition 6.2.1 gives a bijection between the present problem and (j). An elegant bijection with (e) was given by F. Bernhart (private communication, 1996).

j. Let $A(x) = x + x^3 + 2x^4 + 6x^5 + \cdots$ (respectively, $B(x) = x^2 + x^3 + 3x^4 + 8x^5 + \cdots$) be the generating function for Dyck paths from $(0, 0)$ to $(2n, 0)$ $(n > 0)$ such that the path only touches the x-axis at the beginning and end, and the number of steps $(1, -1)$ at the end of the path is odd (respectively, even). Let $C(x) = 1 + x + 2x^2 + 5x^3 + \cdots$ be the generating function for all Dyck paths from $(0, 0)$ to $(2n, 0)$, so the coefficients are Catalan numbers by (i). It is easy to see that $A = x(1 + CB)$ and $B = xCA$. (Also $C = 1/(1 - A - B)$, though we don't need that fact here.) Solving for A gives $A = x/(1 - x^2C^2)$. The generating function we want is $1/(1 - A)$, which simplifies (using $1 + xC^2 = C$) to $1 + xC$, and the proof follows. This result is due to E. Deutsch (private communication, 1996).

k. This result is due to P. Peart and W. Woan, Dyck paths with no peaks at height 2, preprint. The authors give a generating function proof and a simple bijection with (i).

l. The region bounded by the two paths is called a *parallelogram polyomino.* It is an array of unit squares, say with k columns $C_1, \ldots, C_k$. Let a_i be the number of squares in column C_i, for $1 \leq i \leq k$, and let b_i be the number of rows in common to C_i and C_{i+1}, for $1 \leq i \leq k - 1$. Define a sequence σ of 1's and -1's as follows (where exponentiation denotes repetition):

$$\sigma = 1^{a_1}(-1)^{a_1-b_1+1}1^{a_2-b_1+1}(-1)^{a_2-b_2+1}1^{a_3-b_2+1}\cdots 1^{a_k-b_{k-1}+1}(-1)^{a_k}.$$

This sets up a bijection with (r). For the parallelogram polyomino of Figure 6-12 we have $(a_1, \ldots, a_7) = (3, 3, 4, 4, 2, 1, 2)$ and $(b_1, \ldots, b_6) =$

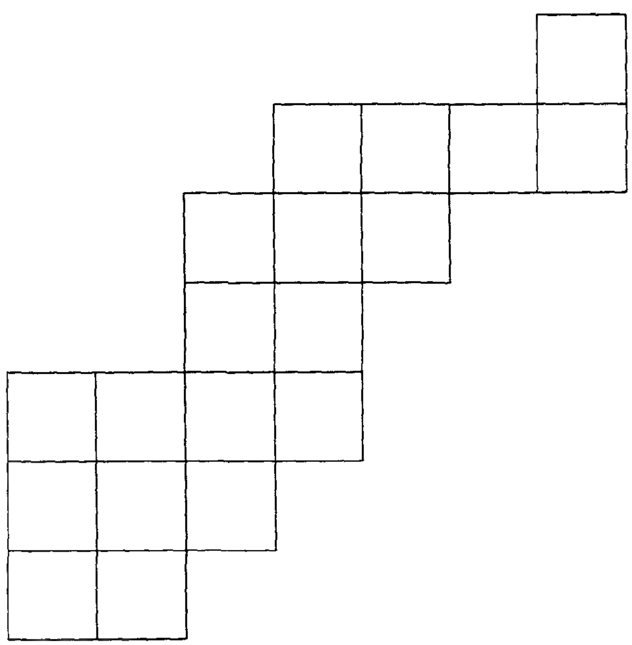


Figure 6-12. A parallelogram polyomino.

(3, 2, 3, 2, 1, 1). Hence (writing $-$ for -1)

$$\sigma = 111-1--111--11---1--1-11--.$$

The enumeration of parallelogram polyominoes is due to J. Levine, *Scripta Math.* **24** (1959), 335–338, and later G. Pólya, *J. Combinatorial Theory* **6** (1969) 102–105. See also L. W. Shapiro, *Discrete Math.* **14** (1976), 83–90; W.-J. Woan, L. W. Shapiro, and D. G. Rogers, *Amer. Math. Monthly* **104** (1997), 926–931; G. Louchard, *Random Structures and Algorithms* **11** (1997), 151–178; and R. A. Sulanke, *J. Difference Equations and Applications*, to appear. For more information on the fascinating topic of polyomino enumeration, see M.-P. Delest and G. Viennot, *Theoretical Computer Science* **34** (1984), 169–206, and X. G. Viennot, in *Séries Formelles et Combinatoire Algébrique* (P. Leroux and C. Reutenauer, eds.), Publications de Laboratoire de Combinatoire et d'Informatique Mathématique **11**, Université du Québec à Montréal, 1992, pp. 399–420.

m. Regarding a path as a sequence of steps, remove the first and last steps from the two paths in (l). This variation of (l) was suggested by L. W. Shapiro (private communication, 1998).

n. Fix a vertex v. Starting clockwise from v, at each vertex write 1 if encountering a chord for the first time and -1 otherwise. This gives a bijection with (r). This result is apparently due to A. Errera, *Mém. Acad. Roy. Belgique Coll. 8º (2)* **11** (1931), 26 pp. See also J. Riordan, *Math. Comp.* **29** (1975), 215–222, and S. Dulucq and J.-G. Penaud, *Discrete Math.* **17** (1993), 89–105.

o. Cut the circle in (n) between two fixed vertices and "straighten out."

p,q. These results are due to I. M. Gelfand, M. I. Graev, and A. Postnikov, in *The Arnold-Gelfand Mathematical Seminars*, Birkhäuser, Boston, 1997, pp. 205–221 (§6). For (p), note that there is always an arc from the leftmost to the rightmost vertex. When this arc is removed, we obtain two smaller trees satisfying the conditions of the problem. This leads to an easy bijection with (c). The trees of (p) are called *noncrossing alternating trees*.

An equivalent way of stating the above bijection is as follows. Let T be a noncrossing alternating tree on the vertex set $1, 2, \dots, n+1$ (in that order from left to right). Suppose that vertex i has r_i neighbors that are larger than i. Let u_i be the word in the alphabet $\{1, -1\}$ consisting of r_i 1's followed by a -1. Let $u(T) = u_1u_2\cdots u_{n+1}$. Then u is a bijection between the objects counted by (p) and (r). It was shown by M. Schlosser that exactly the same definition of u gives a bijection between (q) and (r)! The proof, however, is considerably more difficult than in the case of (p). (A more complicated bijection was given earlier by C. Krattenthaler.)

For further information on trees satisfying conditions (α), (β), and (δ) (called *alternating trees*), see Exercise 5.41.

s. Consider a lattice path P of the type (h). Let a_i be the area above the x-axis between $x = i-1$ and $x = i$, and below P. This sets up a bijection as desired.

t. Subtract $i-1$ from a_i and append a one at the beginning to get (s). This result is closely related to Exercise 6.25(c). If we replace the alphabet $1, 2, \dots, 2(n-1)$ with the alphabet $\overline{n-1}, n-1, \overline{n-2}, n-2, \dots, \overline{1}, 1$ (in that order) and write the new sequence $b_1, b_2, \dots, b_{n-1}$ in reverse order in a column, then we obtain the arrays of R. King, in *Lecture Notes in Physics*,

vol. 50, Springer-Verlag, Berlin/Heidelberg/New York, 1975, pp. 490–499 (see also S. Sundaram, *J. Combinatorial Theory (A)* **53** (1990), 209–238 (Def. 1.1)) that index the weights of the $(n-1)$-st fundamental representation of $\mathrm{Sp}(2(n-1), \mathbb{C})$.

u. Let $b_i = a_i - a_{i+1} + 1$. Replace a_i with one 1 followed by b_i -1's for $1 \le i \le n$ (with $a_{n+1} = 0$) to get (r).

v. Take the first differences of the sequences in (u).

w. Do a depth-first search through a plane tree with $n+1$ vertices as in (e). When a vertex is encountered for the first time, record one less than its number of successors, except that the last vertex is ignored. This gives a bijection with (e).

x. These sequences are just the inversion tables (as defined in Section 1.3) of the 321-avoiding permutations of (ee). For a proof see S. C. Billey, W. Jockusch, and R. Stanley, *J. Alg. Combinatorics* **2** (1993), 345–374 (Thm. 2.1). (The previous reference deals with the *code* $c(w)$ of a permutation w rather than the inversion table $I(w)$. They are related by $c(w) = I(w^{-1})$. Since w is 321-avoiding if and only if w^{-1} is 321-avoiding, it makes no difference whether we work with the code or with the inversion table.)

y. If we replace a_i by $n - a_i$, then the resulting sequences are just the inversion tables of 213-avoiding permutations w (i.e., there does not exist $i < j < k$ such that $w_j < w_i < w_k$). Such permutations are in obvious bijection with the 312-avoiding permutations of (ff). For further aspects of this exercise, see Exercise 6.32.

z. Given a sequence $a_1, \ldots, a_n$ of the type being counted, define recursively a binary tree $T(a_1, \ldots, a_n)$ as follows. Set $T(\emptyset) = \emptyset$. If $n > 0$, then let the left subtree of the root of $T(a_1, \ldots, a_n)$ be $T(a_1, a_2, \ldots, a_{n-a_n})$ and the right subtree of the root be $T(a_{n-a_n+1}, a_{n-a_n+2}, \ldots, a_{n-1})$. This sets up a bijection with (c). Alternatively, the sequences $a_n - 1, a_{n-1} - 1, \ldots, a_1 - 1$ are just the inversion tables of the 312-avoiding permutations of (ff). Let us also note that the sequences $a_1, a_2, \ldots, a_n$ are precisely the sequences $\tau(u)$, $u \in \mathfrak{S}_n$, of Exercise 5.49(d).

aa. If $a = a_1 a_2 \cdots a_k$ is a word in the alphabet $[n-1]$, then let $w(a) = s_{a_1} s_{a_2} \cdots s_{a_k} \in \mathfrak{S}_n$, where s_i denotes the adjacent transposition $(i, i+1)$. Then $w(a) = w(b)$ if $a \sim b$; and the permutations $w(a)$, as a ranges over a set of representatives for the classes B being counted, are just those enumerated by (ee). This statement follows from S. C. Billey, W. Jockusch, and R. Stanley, *J. Alg. Combinatorics* **2** (1993), 345–374 (Thm. 2.1).

bb. Regard a partition whose diagram fits in an $(n-1) \times (n-1)$ square as an order ideal of the poset $(\mathbf{n-1}) \times (\mathbf{n-1})$ in an obvious way. Then the partitions being counted correspond to the order ideals of (ccc). Bijections with other Catalan families were given by D. E. Knuth and A. Postnikov. Postnikov's bijection is the following. Let λ be a partition whose diagram is contained in an $(n-1) \times (n-1)$ square S. Let x be the lower right corner of the Durfee square of λ. Let L_1 be the lattice path from the upper right corner of S to x that follows the boundary of λ. Similarly let L_2 be the lattice path from the lower left corner of S to x that follows the boundary of λ. Reflect L_2 about the main diagonal of S. The paths L_1 and the reflection of L_2 form a pair of paths as in (m). Figure 6-13 illustrates this bijection for $n = 8$ and $\lambda = (5, 5, 3, 3, 3, 1)$. The path L_1 is shown in dark lines and L_2 and its reflection in dashed lines.

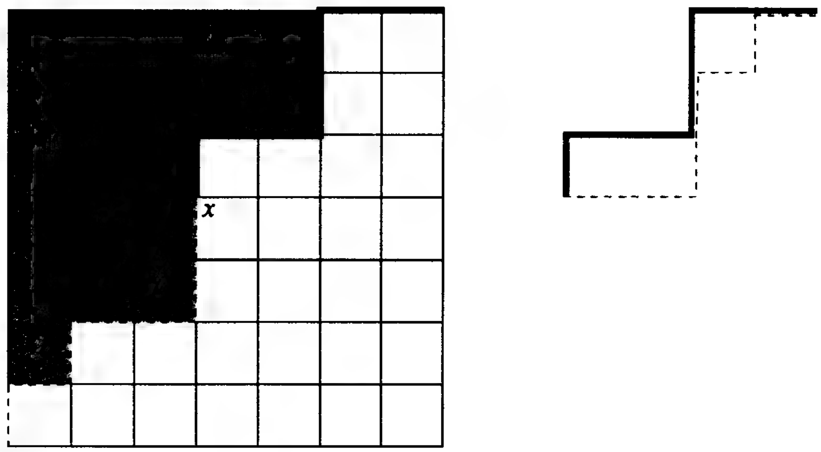


Figure 6-13. A bijection between (ccc) and (bb).

cc. Remove the first occurrence of each number. What remains is a permutation w of $[n]$ that uniquely determines the original sequence. These permutations are precisely the ones in (ff). There is also an obvious bijection between the sequences being counted and the nonintersecting arcs of (o).

dd. Replace each odd number by 1 and even number by -1 to get a bijection with (r).

ee. Corollary 7.23.11 shows that the RSK algorithm (Section 7.11) establishes a bijection with (xx). See also D. E. Knuth, *The Art of Computer Programming*, vol. 3, *Sorting and Searching*, Addison-Wesley, Reading, Massachusetts, 1973 (p. 64).

The earliest explicit enumeration of 321-avoiding permutations seems to be due to J. M. Hammersley, in *Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, University of California Press, Berkeley/Los Angeles, 1972, pp. 345–394. In equation (15.9) he states the result, saying "and this can be proved in general." The first published proof is a combinatorial proof due to D. G. Rogers, *Discrete Math.* **22** (1978), 35–40. Another direct combinatorial proof, based on an idea of Goodman, de la Harpe, and Jones, appears in S. C. Billey, W. Jockusch, and R. Stanley, *J. Alg. Combinatorics* **2** (1993), 345–374 (after the proof of Theorem 2.1). A sketch of this proof goes as follows. Given the 321-avoiding permutation $w = a_1a_2\cdots a_n$, define $c_i = \#\{j : j > i, w_j < w_i\}$. Let $\{j_1, \ldots, j_\ell\}_< = \{j : c_j > 0\}$. Define a lattice path from $(0, 0)$ to (n, n) as follows. Walk horizontally from $(0, 0)$ to $(c_{j_1} + j_1 - 1, 0)$, then vertically to $(c_{j_1} + j_1 - 1, j_1)$, then horizontally to $(c_{j_2} + j_2 - 1, j_1)$, then vertically to $(c_{j_2} + j_2 - 1, j_2)$, etc. The last part of the path is a vertical line from $(c_{j_\ell} + j_\ell - 1, j_{\ell-1})$ to $(c_{j_\ell} + j_\ell - 1, j_\ell)$, then (if needed) a horizontal line to (n, j_ℓ), and finally a vertical line to (n, n). This establishes a bijection with (h).

For an elegant bijection with (ff), see R. Simion and F. W. Schmidt, *Europ. J. Combinatorics* **6** (1985), 383–406 (Prop. 19). Two other bijections with (ff) appear in D. Richards, *Ars Combinatoria* **25** (1988), 83–86, and J. West, *Discrete Math.* **146** (1995) 247-262 (Thm. 2.8).

ff. There is an obvious bijection between 312-avoiding and 231-avoiding permutations, viz., $a_1a_2\cdots a_n \mapsto n+1-a_n, \ldots, n+1-a_2, n+1-a_1$. It is easily seen that the 231-avoiding permutations are the same as those of (ii), as first observed by D. E. Knuth [**5.41**, Exer. 2.2.1.5]. The enumeration *via* Catalan numbers appears in *ibid.*, Exer. 2.2.1.4. References to bijections with (ee) are given in the solution to (ee). For the problem of counting permutations in $\mathfrak{S}_n$ according to the number of subsequences with the pattern 132 (equivalently, 213, 231, or 312), see M. Bóna, in *Conference Proceedings*, vol. 1, *Formal Power Series and Algebraic Combinatorics, July 14–July 18, 1997,* Universität Wien, pp. 107–118.

gg. This result appears on p. 796 of D. M. Jackson, *Trans. Amer. Math. Soc.* **299** (1987), 785–801, but probably goes back much earlier. For a direct bijective proof, it is not hard to show that the involutions counted here are the same as those in (kk).

hh. A coding of planar maps due to R. Cori, *Astérisque* **27** (1975), 169 pp., when restricted to plane trees, sets up a bijection with (e).

ii. When an element a_i is put on the stack, record a 1. When it is taken off, record a -1. This sets up a bijection with (r). This result is due to D. E. Knuth [**5.41**, Exer. 2.2.1.4]. The permutations being counted are just the 231-avoiding permutations, which are in obvious bijection with the 312-avoiding permutations of (ff) (see Knuth, *ibid.*, Exer. 2.2.1.5).

jj. Same set as (ee), as first observed by R. Tarjan, *J. Assoc. Comput. Mach.* **19** (1972), 341–346 (the case $m = 2$ of Lemma 2). The concept of queue sorting is due to Knuth [**5.41**, Ch. 2.2.1].

kk. Obvious bijection with (o).

ll. See I. M. Gessel and C. Reutenauer, *J. Combinatorial Theory (A)* **64** (1993), 189–215 (Thm. 9.4 and discussion following).

mm. This result is due to O. Guibert and S. Linusson, in *Conference Proceedings*, vol. 2, *Formal Power Series and Algebraic Combinatorics, July 14 – July 18, 1997*, Universität Wien, pp. 243–252. Is there a nice formula for the number of alternating Baxter permutations of $[m]$?

nn. This is the same set as (ee). See Theorem 6.2.1 of the reference given in (ee) to S. C. Billey *et al.* For a generalization to other Coxeter groups, see J. R. Stembridge, *J. Alg. Combinatorics* **5** (1996), 353–385.

oo. These are just the 132-avoiding permutations $w_1\cdots w_n$ of $[n]$ (i.e., there does not exist $i < j < k$ such that $w_i < w_k < w_j$), which are in obvious bijection with the 312-avoiding permutations of (ff). This result is an immediate consequence of the following results: (i) I. G. Macdonald, *Notes on Schubert Polynomials*, Publications du LACIM, vol. 6, Univ. du Québec à Montréal, 1991, (4.7) and its converse stated on p. 46 (due to A. Lascoux and M. P. Schützenberger), (ii) *ibid.*, eqn. (6.11) (due to Macdonald), (iii) part (ff) of this exercise, and (iv) the easy characterization of dominant permutations (as defined in Macdonald, *ibid.*) as 132-avoiding permutations. For a simpler proof of the crucial (6.11) of Macdonald, see S. Fomin and R. Stanley, *Advances in Math.* **103** (1994), 196–207 (Lemma 2.3).

pp. See Exercise 3.68(b). Noncrossing partitions first arose in the work of H. W. Becker, *Math. Mag.* **22** (1948–49), 23–26, in the form of *planar rhyme*

schemes, i.e., rhyme schemes with no crossings in the *Puttenham diagram*, defined by G. Puttenham, *The Arte of English Poesie*, London, 1589 (pp. 86–88). Further results on noncrossing partitions are given by H. Prodinger, *Discrete Math.* **46** (1983), 205–206; N. Dershowitz and S. Zaks, *Discrete Math.* **62** (1986), 215–218; R. Simion and D. Ullman, *Discrete Math.* **98** (1991), 193–206; P. H. Edelman and R. Simion, *Discrete Math.* **126** (1994), 107–119; R. Simion, *J. Combinatorial Theory (A)* **65** (1994), 270–301; R. Speicher, *Math. Ann.* **298** (1994), 611–628; A. Nica and R. Speicher, *J. Algebraic Combinatorics* **6** (1997), 141–160; R. Stanley, *Electron. J. Combinatorics* **4**, R20 (1997), 14 pp. See also Exercise 5.35.

qq. These partitions are clearly the same as the noncrossing partitions of (pp). This description of noncrossing partitions is due to R. Steinberg (private communication).

rr. Obvious bijection with (pp). (Vertical lines are in the same block if they are connected by a horizontal line.) As mentioned in the Notes to Chapter 1, Murasaki diagrams were used in *The Tale of Genji* to represent the 52 partitions of a five-element set. The noncrossing Murasaki diagrams correspond exactly to the noncrossing partitions. The statement that noncrossing Murasaki diagrams are enumerated by Catalan numbers seems first to have been observed by H. W. Gould, who pointed it out to M. Gardner, leading to its mention in [**27**]. Murasaki diagrams were not actually used by Lady Murasaki herself. It wasn't until the Wasan period of old Japanese mathematics, from the late 1600s well into the 1700s, that the Wasanists started attaching the Murasaki diagrams (which were actually incense diagrams) to illustrated editions of *The Tale of Genji*.

ss. This result was proved by M. Klazar, *Europ. J. Combinatorics* **17** (1996), 53–68 (p. 56), using generating function techniques.

tt. See R. C. Mullin and R. G. Stanton, *Pacific J. Math.* **40** (1972), 167–172 (p. 168). They set up a bijection with (e). They also show that $2n+1$ is the largest possible value of k for which there exists a noncrossing partition of $[k]$ with $n+1$ blocks such that no block contains two consecutive integers. A simple bijection with (a) was given by D. P. Roselle, *Utilitas Math.* **6** (1974), 91–93. The following bijection with (d) is due to A. Vetta (1997). Label the vertices $1, 2, \ldots, 2n+1$ of a tree in (d) in preorder. Define i and j to be in the same block of $\pi \in \Pi_{2n+1}$ if j is a right child of i.

uu. Let P_n denote the poset of intervals with at least two elements of the chain $\boldsymbol{n}$, ordered by inclusion. Let $\mathcal{A}_n$ denote the set of antichains of P_n. By the last paragraph of Section 3.1, $\#\mathcal{A}_n$ is equal to the number of order ideals of P_n. But P_n is isomorphic to the poset $\operatorname{Int}(\boldsymbol{n-1})$ of *all* (nonempty) intervals of $\boldsymbol{n-1}$, so by (bbb) we have $\#\mathcal{A}_n = C_n$. Given an antichain $A \in \mathcal{A}_n$, define a partition π of $[n]$ by the condition that i and j (with $i < j$) belong to the same block of π if $[i, j] \in A$ (and no other conditions not implied by these). This establishes a bijection between $\mathcal{A}_n$ and the nonnesting partitions of $[n]$. For a further result on nonnesting partitions, see the solution to Exercise 5.44. The present exercise was obtained in collaboration with A. Postnikov. The concept of nonnesting partitions for any reflection group (with the present case corresponding to the symmetric group $\mathfrak{S}_n$) is due to

Postnikov and is further developed in C. A. Athanasiadis, On noncrossing and nonnesting partitions for classical reflection groups, preprint, 1998.

vv. If $\lambda = (\lambda_1, \ldots, \lambda_{n-1}) \subseteq (n-1, n-2, \ldots, 1)$, then the sequences $(1, \lambda_{n-1} + 1, \ldots, \lambda_1 + 1)$ are in bijection with (s). Note also that the set of Young diagrams contained in $(n-1, n-2, \ldots, 1)$, ordered by inclusion (i.e., the interval $[\emptyset, (n-1, n-2, \ldots, 0)]$ in Young's lattice, as defined in Exercise 3.63), is isomorphic to $J(\mathrm{Int}(\boldsymbol{n}-\mathbf{1}))^*$, thereby setting up a bijection with (bbb).

ww. Given a standard Young tableau T of shape (n, n), define $a_1 a_2 \cdots a_{2n}$ by $a_i = 1$ if i appears in row 1 of T, while $a_i = -1$ if i appears in row 2. This sets up a bijection with (r). See also [**7.72**, p. 63] and our Proposition 7.10.3.

xx. See the solution to (ee) (first paragraph) for a bijection with 321-avoiding permutations. An elegant bijection with (ww) appears in [**2.15**, vol. 1, p. 131] (repeated in [**7.72**, p. 63] Namely, given a standard Young tableau T of shape (n, n), let P consist of the part of T containing the entries $1, 2, \ldots, n$; while Q consists of the complement in T of P, rotated 180°, with the entry i replaced by $2n + 1 - i$. See also Corollary 7.23.12.

yy. Let b_i be the number of entries in row i that are equal to $n - i + 1$ (so $b_n = 0$). The sequences $b_n + 1, b_{n-1} + 1, \ldots, b_1 + 1$ obtained in this way are in bijection with (s).

zz. This result is equivalent to Prop. 2.1 of S. C. Billey, W. Jockusch, and R. Stanley, *J. Algebraic Combinatorics* **2** (1993), 345–374. See also the last paragraph on p. 363 of this reference.

aaa. Obvious bijection with (ww). This interpretation of Catalan numbers appears in [**3.34**, p. 222]. Note also that if we label the elements of $\mathbf{2} \times \boldsymbol{n}$ analogously to what was illustrated for $n = 3$, then the linear extensions coincide with the permutations of (dd).

bbb. There is an obvious bijection with order ideals I of $\mathrm{Int}(\boldsymbol{n})$ that contain every one-element interval of $\boldsymbol{n}$. But the "upper boundary" of the Hasse diagram of I "looks like" the Dyck paths of (i). See [**3.34**, bottom of p. 222].

ccc. This result is equivalent to the $q = 1$ case of G. E. Andrews, *J. Statist. Plann. Inf.* **34** (1993), 19–22 (Corollary 1). For a more explicit statement and some generalizations, see R. G. Donnelly, Ph.D. thesis, University of North Carolina, 1997, and Symplectic and odd orthogonal analogues of $L(m, n)$, preprint. For a bijective proof, see the solution to (bb). A sequence of posets interpolating between the poset $\mathrm{Int}(\boldsymbol{n}-\mathbf{1})$ of (bbb) and A_{n-1}, and all having C_n order ideals, was given by D. E. Knuth (private communication, 9 December 1997).

ddd. Given a sequence $1 \leq a_1 \leq \cdots \leq a_n$ of integers with $a_i \leq i$, define a poset P on the set $\{x_1, \ldots, x_n\}$ by the condition that $x_i < x_j$ if and only if $j + a_{n+1-i} \geq n + 1$. (Equivalently, if Z is the matrix of the zeta function of P, then the 1's in $Z - I$ form the shape of the Young diagram of a partition, rotated 90° clockwise and justified into the upper right-hand corner.) This yields a bijection with (s). This result is due to R. L. Wine and J. E. Freund, *Ann. Math. Statist.* **28** (1957), 256–259. See also R. A. Dean and G. Keller, *Canad. J. Math.* **20** (1968), 535–554. Such posets are now called *semiorders*. For further information, see P. C. Fishburn, *Interval Orders and Interval Graphs*, Wiley-Interscience, New York,

1985, and W. T. Trotter, *Combinatorics and Partially Ordered Sets*, Johns Hopkins University Press, Baltimore/London, 1992 (Ch. 8). For the labeled version of this exercise, see Exercise 6.30.

eee. The lattice $J(P)$ of order ideals of the poset P has a natural planar Hasse diagram. There will be two elements covering $\hat{0}$, corresponding to the two minimal elements of P. Draw the Hasse diagram of $J(P)$ so that the rooted minimal element of P goes to the left of $\hat{0}$ (so the other minimal element goes to the right). The "outside boundary" of the Hasse diagram then "looks like" the pair of paths in (l) (rotated 45° counterclockwise).

fff. These relations are called *similarity relations*. See L. W. Shapiro, *Discrete Math.* **14** (1976), 83–90; V. Strehl, *Discrete Math.* **19** (1977), 99–101; D. G. Rogers, *J. Combinatorial Theory (A)* **23** (1977), 88–98; J. W. Moon, *Discrete Math.* **26** (1979), 251–260. Moon gives a bijection with (r). E. Deutsch (private communication) has pointed out an elegant bijection with (h), viz., the set enclosed by a path and its reflection in the diagonal *is* a similarity relation (as a subset of $[n] \times [n]$). The connectedness of the columns ensures the last requirement in the definition of a similarity relation.

ggg. A simple combinatorial proof was given by L. W. Shapiro, *J. Combinatorial Theory* **20** (1976), 375–376. Shapiro observes that this result is a combinatorial manifestation of the identity

$$\sum_{k\geq 0} \binom{n}{2k} 2^{n-2k} C_k = C_{n+1},$$

due to J. Touchard, in *Proc. Int. Math. Congress, Toronto (1924)*, vol. 1, 1928 (p. 465).

hhh. Obvious bijection with (bbb). This interpretation in terms of stacking coins is due to J. Propp. See A. M. Odlyzko and H. S. Wilf, *Amer. Math. Monthly* **95** (1988), 840–843 (Rmk. 1).

* **iii.** See J. H. van Lint, *Combinatorial Theory Seminar, Eindhoven University of Technology*, Lecture Notes in Mathematics, **382**, Springer-Verlag, Berlin/Heidelberg/New York, 1974 (pp. 22 and 26–27).

jjj. The total number of n-element multisets on $\mathbb{Z}/(n+1)\mathbb{Z}$ is $\binom{2n}{n}$ (see Section 1.2). Call two such multisets M and N *equivalent* if for some $k \in \mathbb{Z}/(n+1)\mathbb{Z}$ we have $M = \{a_1, \ldots, a_n\}$ and $N = \{a_1+k, \ldots, a_n+k\}$. This defines an equivalence relation in which each equivalence class contains $n+1$ elements, exactly one of which has its elements summing to 0. Hence the number of multisets with elements summing to 0 (or to any other fixed element of $\mathbb{Z}/(n+1)\mathbb{Z}$) is $\frac{1}{n+1}\binom{2n}{n}$. This result appears in R. K. Guy, *Amer. Math. Monthly* **100** (1993), 287–289 (with a more complicated proof due to I. Gessel).

kkk. This result is implicit in the paper G. X. Viennot, *Astérisque* **121–122** (1985), 225–246. Specifically, the bijection used to prove (12), when restricted to Dyck words, gives the desired bijection. A simpler bijection follows from the work of J.-G. Penaud, in *Séminaire Lotharingien de Combinatoire*, 22^e Session, Université Louis Pasteur, Strasbourg, 1990, pp. 93–130 (Cor. IV-2-8). Yet another proof follows from more general results of J. Bétréma and J.-G. Penaud, *Theoret. Comput. Sci.* **117** (1993), 67–88. For some related problems, see Exercise 6.46.

* **lll.** Let I be an order ideal of the poset Int $(n-1)$ defined in (bbb). Associate with I the set R_I of all points $(x_1, \ldots, x_n) \in \mathbb{R}^n$ satisfying $x_1 > \cdots > x_n$ and $x_i - x_j < 1$ if $[i, j-1] \in I$. This sets up a bijection between (bbb) and the regions R_I being counted. This result is implicit in R. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **93** (1996), 2620–2625 (§2), and also appears (as part of more general results) in C. A. Athanasiadis, Ph.D. thesis, Massachusetts Institute of Technology, 1996 (Cor. 7.1.3) and A. Postnikov and R. Stanley, Deformations of Coxeter hyperplane arrangements, preprint (Prop. 7.2), available at http://front.math.ucdavis.edu/math.CO/9712213.

* **mmm.** Let P be a convex $(n+2)$-gon with vertices $v_1, v_2, \ldots, v_{n+2}$ in clockwise order. Let T be a triangulation of T as in (a), and let a_i be the number of triangles incident to v_i. Then the map $T \mapsto (a_1, \ldots, a_{n+2})$ establishes a bijection with (a). This remarkable result is due to J. H. Conway and H. S. M. Coxeter [**20**, problems (28) and (29)]. The arrays (6.54) are called *frieze patterns*.

nnn. See F. T. Leighton and M. Newman, *Proc. Amer. Math. Soc.* **79** (1980), 177–180, and L. W. Shapiro, *Proc. Amer. Math. Soc.* **90** (1984), 488–496.

6.20. a. Given a path P of the first type, let (i, i) be the first point on P that intersects $y = x$. Replace the portion of P from $(1, 0)$ to (i, i) by its reflection about $y = x$. This yields the desired bijection.

This argument is the famous "reflection principle" of D. André, *C. R. Acad. Sci. Paris* **105** (1887), 436–437. The application (b) below is also due to André. The importance of the reflection principle in combinatorics and probability theory was realized by W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1, John Wiley and Sons, New York, 1950 (3rd edition, 1968). For a vast number of extensions and ramifications, see L. Takács, *Combinatorial Methods in the Theory of Stochastic Processes*, John Wiley and Sons, New York, 1967; T. V. Narayana, *Lattice Path Combinatorics with Statistical Applications*, Mathematical Expositions no. 23, University of Toronto Press, Toronto, 1979; and S. G. Mohanty, *Lattice Path Counting and Applications*, Academic Press, New York, 1979. For a profound generalization of the reflection principle based on the theory of Coxeter groups, as well as some additional references, see I. Gessel and D. Zeilberger, *Proc. Amer. Math. Soc.* **115** (1992), 27–31.

b. The first step in such a lattice path must be from $(0, 0)$ to $(1, 0)$. Hence we must subtract from the total number of paths from $(1, 0)$ to (m, n) the number that intersect $y = x$, so by (a) we get $\binom{m+n-1}{n} - \binom{m+n-1}{m} = \frac{m-n}{m+n}\binom{m+n}{n}$.

c. Move the path one unit to the right to obtain the case $m = n + 1$ of (b).

6.21. a. Given a path $P \in X_n$, define $c(P) = (c_0, c_1, \ldots, c_n)$, where c_i is the number of horizontal steps of P at height $y = i$. It is not difficult to verify that the cyclic permutations $C_j = (c_j, c_{j+1}, \ldots, c_n, c_1, \ldots, c_{j-1})$ of $c(P)$ are all distinct, and for each such there is a unique $P_j \in X_n$ with $c(P_j) = C_j$. Moreover, the number of excedances of the paths $P = P_0, P_1, \ldots, P_n$ are just the numbers $0, 1, \ldots, n$ in some order. From these observations the proof is immediate.

This result, known as the *Chung–Feller theorem*, is due to K. L. Chung and W. Feller, *Proc. Nat. Acad. Sci. U.S.A.* **35** (1949), 605–608. A refinement was given by T. V. Narayana, *Skand. Aktuarietidskr.* **1967** (1967), 23–30. For

further information, see the books by Narayana (§I.2) and Mohanty (§3.3) mentioned in the solution to Exercise 6.20(a).

b. Immediate from (a).

6.22. This result was given by L. W. Shapiro, problem E2903, *Amer. Math. Monthly* **88** (1981), 619. An incorrect solution appeared in **90** (1983), 483–484. A correct but nonbijective solution was given by D. M. Bloom, **92** (1985), 430. The editors asked for a bijective proof in problem E3096, **92** (1985), 428, and such a proof was given by W. Nichols, **94** (1987), 465–466. Nichols's bijection is the following. Regard a lattice path as a sequence of E's (for the step $(1,0)$) and N's (for the step $(0, 1)$). Given a path P of the type we are enumerating, define recursively a new path $\psi(P)$ as follows:

$$\psi(\emptyset) = \emptyset, \qquad \psi(DX) = D\psi(X), \qquad \psi(D'X) = E\psi(X)ND^*,$$

where (a) D is a path of positive length, with endpoints on the diagonal $x = y$ and all other points below the diagonal, (b) D' denotes the path obtained from D by interchanging E's and N's, and (c) $D = ED^*N$. Then ψ establishes a bijection between the paths we are enumerating and the paths of Exercise 6.19(h) with n replaced by $2n$. For an explicit description of ψ^{-1} and a proof that ψ is indeed a bijection, see the solution of Nichols cited above.

6.23. The Black pawn on a6 must promote to a knight and then move (in a unique way) to h7 in five additional moves. The Black pawn on a7 must also promote to a knight and then move (in a unique way) to f8 in four additional moves. White then plays Pf7 mate. The first move must be Pa5, after which the number of solutions is the same as if the pawn on a7 were on a6. Each pawn then makes nine moves (including moves after promotion). After the first move Pa5, denote a move by the pawn on a5 by $+1$ and a move by the pawn on a7 by -1. Since the pawn on a7 can never overtake the pawn on a5 (even after promotion), it follows that the number of solutions is just the number of sequences of nine 1's and nine -1's with all partial sums nonnegative. By Exercise 6.19(r), the number of solutions is therefore the Catalan number $C_9 = 4862$.

This problem is due to Kauko Väisänen, and appears in A. Puusa, *Queue Problems*, Finnish Chess Problem Society, Helsinki, 1992 (Problem 2). This booklet contains fifteen problems of a similar nature. See also Exercise 7.18. For more information on serieshelpmates in general, see A. Dickins, *A Guide to Fairy Chess*, Dover, New York, 1971, p. 10, and J. M. Rice and A. Dickins, *The Serieshelpmate*, second edition, Q Press, Kew Gardens, 1978.

6.24. These are just Catalan numbers! See for instance J. Gili, *Catalan Grammar*, Dolphin, Oxford, 1993, p. 39. A related question appears in *Amer. Math. Monthly* **103** (1996), 538 and 577.

6.25. **a.** Follows from Exercise 3.29(b) and 6.19(bbb). See L. W. Shapiro, *American Math. Monthly* **82** (1975), 634–637.

b. We assume knowledge of Chapter 7. It follows from the results of Appendix 2 of Chapter 7 that we want the coefficient of the trivial Schur function $s_\emptyset$ in the Schur function expansion of $(x_1+x_2)^{2n}$ in the ring $\Xi_2 = \Lambda_2/(x_1x_2 - 1)$. Since $s_\emptyset = s_{(n,n)}$ in Ξ_2, the number we want is just $\langle s_1^{2n}, s_{(n,n)}\rangle = f^{(n,n)}$ (using Corollary 7.12.5), and the result follows from Exercise 6.19(ww).

c. See R. Stanley, *Ann. New York Acad. Sci.*, vol. 576, 1989, pp. 500–535 (Example 4 on p. 523).

d. See R. Stanley, in *Advanced Studies in Pure Math.*, vol. 11, Kinokuniya, Tokyo, and North-Holland, Amsterdam/New York, 1987, pp. 187–213 (bottom of p. 194). A simpler proof follows from R. Stanley, *J. Amer. Math. Soc.* **5** (1992), 805–851 (Prop. 8.6). For a related result, see C. Chan, *SIAM J. Disc. Math.* **4** (1991), 568–574.

e. See L. R. Goldberg, *Adv. Math.* **85** (1991), 129–144 (Thm. 1.7).

f. See D. Tischler, *J. Complexity* **5** (1989), 438–456.

g. This algebra is the *Temperley–Lieb algebra* $A_{\beta,n}$ (over K), with many interesting combinatorial properties. For its basic structure see F. M. Goodman, P. de la Harpe, and V. F. R. Jones, *Coxeter Graphs and Towers of Algebras*, Springer-Verlag, New York, 1989, p. 33 and §2.8. For a direct connection with 321-avoiding permutations (defined in Exercise 6.19(ee)), see S. C. Billey, W. Jockusch, and R. Stanley, *J. Algebraic Combinatorics* **2** (1993), 345–374 (pp. 360–361).

h. See J.-Y. Shi, *Quart. J. Math.* **48** (1997), 93–105 (Thm. 3.2(a)).

i. This remarkable conjecture is a small part of a vast conjectured edifice due to M. Haiman, *J. Algebraic Combinatorics* **3** (1994), 17–76. See also A. M. Garsia and M. Haiman, *J. Algebraic Combinatorics* **5** (1996), 191–244; A. M. Garsia and M. Haiman, *Electron. J. Combinatorics* **3** (1996), no. 2, Paper 24; and M. Haiman, (t, q)-Catalan numbers and the Hilbert scheme, *Discrete Math.*, to appear.

6.26. a. This curious result can be proved by induction using suitable row and column operations. It arose from a problem posed by E. Berlekamp and was solved by L. Carlitz, D. P. Roselle, and R. A. Scoville, *J. Combinatorial Theory* **11** (1971), 258–271. A slightly different way of stating the result appears in [**3.34**, p. 223].

b. *Answer:* $a_n = C_n$, the nth Catalan number. One way (of many) to prove this result is to apply part (a) to the cases $\lambda = (2n+1, 2n, \ldots, 2, 1)$ and $\lambda = (2n, 2n-1, \ldots, 2, 1)$, and to use the interpretation of Catalan numbers given by Corollary 6.2.3(v). Related work appears in A. Kellogg (proposer), Problem 10585, *Amer. Math. Monthly* **104** (1997), 361, and C. Radoux, *Bull. Belgian Math. Soc. (Simon Stevin)* **4** (1997), 289–292.

6.27. a. The unique such basis $y_0, y_1, \ldots, y_n$, up to sign and order, is given by

$$y_j = \sum_{i=0}^{j} (-1)^{j-i} \binom{i+j}{2i} x_i.$$

b. Now

$$y_j = \sum_{i=0}^{j} (-1)^{j-i} \binom{i+j+1}{2i+1} x_i.$$

6.28. a. The problem of computing the probability of convexity was raised by J. van de Lune and solved by R. B. Eggleton and R. K. Guy, *Math. Mag.* **61** (1988), 211–219, by a clever integration argument. The proof of Eggleton and Guy can be "combinatorialized" so that integration is avoided. The

decomposition of $\mathcal{C}_d$ given below in the solution to (c) also yields a proof. For a more general result, see P. Valtr, in *Intuitive Geometry (Budapest, 1995)*, Bolyai Soc. Math. Stud. **6**, János Bolyai Math Soc., Budapest, 1997, pp. 441–443.

b. Suppose that $x = (x_1, x_2, \dots, x_d) \in \mathcal{C}_d$. We say that an index i is *slack* if $2 \le i \le d-1$ and $x_{i-1} + x_{i+1} > 2x_i$. If no index is slack, then either $x = (0, 0, \dots, 0)$, $x = (1, 1, \dots, 1)$, or $x = \lambda(1, 1, \dots, 1) + (1-\lambda)y$ for $y \in \mathcal{C}_d$ and sufficiently small $\lambda > 0$. Hence in this last case x is not a vertex. So suppose that x has a slack index. If for all slack indices i we have $x_i = 0$, then x is of the stated form (6.55). Otherwise, let i be a slack index such that $x_i > 0$. Let $j = i - p$ be the largest index such that $j < i$ and j is not slack. Similarly, let $k = i + q$ be the smallest index such that $k > i$ and k is not slack. Let

$$A(\epsilon) = \left(x_1, \dots, x_j, x_{j+1} + \frac{\epsilon}{p}, x_{j+2} + \frac{2\epsilon}{p}, \dots,\right.$$
$$\left. x_i + \epsilon, \dots, x_{k-2} + \frac{2\epsilon}{q}, x_{k-1} + \frac{\epsilon}{q}, x_k, \dots, x_n\right).$$

For small $\epsilon > 0$, both $A(\epsilon)$ and $A(-\epsilon)$ are in $\mathcal{C}_d$. Since $x = \frac{1}{2}\big[A(\epsilon) + \frac{1}{2}A(-\epsilon)\big]$, it follows that x is not a vertex. The main idea of this argument is due to A. Postnikov.

c. For $1 \le r \le s \le d$, let

$$\begin{aligned} F_{rs} &= \{(x_1, \dots, x_d) \in \mathcal{C}_d : x_r = x_{r+1} = \cdots = x_s = 0\} \\ F_r^- &= \{(x_1, \dots, x_d) \in \mathcal{C}_d : x_r = x_{r+1} = \cdots = x_d = 0\} \\ F_s^+ &= \{(x_1, \dots, x_d) \in \mathcal{C}_d : x_1 = x_2 = \cdots = x_s = 0\}. \end{aligned}$$

Now F_r^- is a simplex with vertices $\left(1, \frac{k-1}{k}, \frac{k-2}{k}, \dots, \frac{1}{k}, 0, 0, \dots, 0\right)$ for $1 \le k \le r-1$, together with $(0, 0, \dots, 0)$. These vertices have denominators (i.e., the smallest positive integer whose product with the vertex has integer coordinates) $1, 2, 3, \dots, r-1, 1$, respectively. Hence

$$\sum_{n \ge 0} i(F_r^-, n)x^n = \frac{1}{[1][r-1]!}.$$

Similarly

$$\sum_{n \ge 0} i\left(F_s^+, n\right)x^n = \frac{1}{[1][d-s]!}.$$

Since $F_{rs} \cong F_r^- \times F_s^+$, we have $i(F_{rs}, n) = i(F_r^-, n)i(F_s^+, n)$ and

$$\sum_{n \ge 0} i(F_{rs}, n)x^n = \frac{1}{[1][r]!} * \frac{1}{[1][d-s]!}.$$

Let P be the poset of all F_{rs}'s, ordered by inclusion, and let μ denote the Möbius function of $P \cup \{\hat{1}\}$. Let $G = \bigcup_{r=1}^d F_{rr}$, a polyhedral complex in $\mathbb{R}^d$. By Möbius inversion we have

$$i(G, n) = -\sum_{F_{st} \in P} \mu(F_{st}, \hat{1})i(F_{st}, n).$$

But $F_{tu} \subseteq F_{rs}$ if and only if $t \leq r \leq s \leq u$, from which it is immediate that

$$-\mu(F_{st}, \hat{1}) = \begin{cases} 1, & s = t \\ -1, & s = t-1 \\ 0, & \text{otherwise.} \end{cases}$$

Hence

$$\sum_{n \geq 0} i(G, n)x^n = \sum_{r=1}^{d} \frac{1}{[1][r-1]!} * \frac{1}{[1][d-r]!}$$
$$- \sum_{r=1}^{d-1} \frac{1}{[1][r-1]!} * \frac{1}{[1][d-1-r]!}.$$

Now the entire polytope $\mathcal{C}_d$ is just a cone over G with apex $(1, 1, \ldots, 1)$. From this it is not hard to deduce that

$$\sum_{n \geq 0} i(\mathcal{C}_d, n)x^n = \frac{1}{1-x} \sum_{n \geq 0} i(G, n)x^n,$$

and the proof follows.

6.29. See P. Valtr, *Discrete Comput. Geom.* **13** (1995), 637–643. Valtr also shows in *Combinatorica* **16** (1996), 567–573, that if n points are chosen uniformly and independently from inside a triangle, then the probability that the points are in convex position is $\frac{2^n}{(2n)!}\binom{3n-3}{n-1,n-1,n-1}$.

6.30. Equation (6.57) is equivalent to

$$\sum_{n \geq 0} f_n \frac{1}{n!}[\log(1-x)^{-1}]^n x^n = C(x).$$

Hence by (5.25) we need to show that

$$n!\, C_n = \sum_{k=1}^{n} c(n,k) f_k,$$

where $c(n, k)$ is the number of permutations $w \in \mathfrak{S}_n$ with k cycles. Choose a permutation $w \in \mathfrak{S}_n$ with k cycles in $c(n, k)$ ways. Let the cycles of w be the elements of a semiorder P in f_k ways. For each cycle $(a_1, \ldots, a_i)$ of w, replace this element of P with an antichain whose elements are labeled $a_1, \ldots, a_i$. If $a = (a_1, \ldots, a_i)$ and $b = (b_1, \ldots, b_j)$ are two cycles of w, then define $a_r < b_s$ if and only if $a < b$ in P. In this way we get a poset $\rho(P, w)$ with vertices $1, 2, \ldots, n$. It is not hard to see that $\rho(P, w)$ is a semiorder, and that every isomorphism class of n-element semiorders occurs exactly $n!$ times among the posets $\rho(P, w)$. Since by Exercise 6.19(ddd) there are C_n nonisomorphic n-element semiorders, the proof follows.

This result was first proved by J. L. Chandon, J. Lemaire, and J. Pouget, *Math. Sci. Hum.* **62** (1978), 61–80, 83. For a more general situation in which the number A_n of unlabeled objects is related to the number B_n of labeled objects by $\sum B_n(x^n/n!) = \sum A_n(1-e^{-x})^n$, see R. Stanley, *Proc. Natl. Acad. Sci. U.S.A.* **93** (1996), 2620–2625 (Thm. 2.3) and A. Postnikov and R. Stanley,

Deformations of Coxeter hyperplane arrangements, preprint (§6), available at http://front.math.ucdavis.edu/math.CO/9712213.

6.31. **a.** (Sketch.) We will triangulate $\mathcal{P}$ into d-dimensional simplices σ, all containing 0. Thus each σ will have d vertices of the form $e_i - e_j$, where $i < j$. Given a graph G with d edges on the vertex set $[d+1]$, let σ_G be the convex hull of all vectors $e_i - e_j$ for which ij is an edge of G with $i < j$, and let $\tilde{\sigma}_G$ be the convex hull of σ_G and the origin. It is easy to see that $\tilde{\sigma}_G$ is a d-dimensional simplex if and only if G is a tree. Moreover, it can be shown that σ_G lies on the boundary of $\mathcal{P}$ (and hence can be part of a triangulation of the type we are looking for) if and only if G is an *alternating* tree, as defined in Exercise 5.41. We therefore want to choose a set $\mathcal{T}$ of alternating trees T on $[d+1]$ such that the $\tilde{\sigma}_T$'s are the facets of a triangulation of $\mathcal{P}$. One way to do this is to take $\mathbb{T}$ to consist of the *noncrossing* alternating trees on $[d+1]$, i.e., alternating trees such that if $i<j<k<l$, then not both ik and jl are edges. By Exercise 6.19(p) the number of such trees is C_d. (We can also take $\mathbb{T}$ to consist of alternating trees on $[d+1]$ such that if $i<j<k<l$ then not both il and jk are edges. By Exercise 6.19(q) the number of such trees is again C_d.) Moreover, it is easy to see that for any tree T on $[d+1]$ we have $V(\tilde{\sigma}_T) = 1/d!$, where V denotes relative volume. Hence $V(\mathcal{P}) = C_d/d!$. This result appears in I. M. Gelfand, M. I. Graev and A. Postnikov, Combinatorics of hypergeometric functions associated with positive roots, in *The Arnold-Gelfand Mathematical Seminars,* Birkhäuser, Boston, 1997, pp. 205–221 (Thm. 2.3(2)).

b. Order the $\binom{d+1}{2}$ edges ij, $1 \le i < j \le d+1$, lexicographically, e.g., $12 < 13 < 14 < 23 < 24 < 34$. Order the C_d noncrossing alternating trees $T_1, T_2, \ldots, T_{C_d}$ lexicographically by edge set, i.e., $T_i < T_j$ if for some k the first (in lexicographic order) k edges of T_i and T_j coincide, while the $(k+1)$-st edge of T_i precedes the $(k+1)$-st edge of T_j. For instance, when $d=3$ the ordering on the noncrossing alternating trees (denoted by their set of edges) is

$$\{12, 13, 14\},\ \{12, 14, 34\},\ \{13, 14, 23\},\ \{14, 23, 24\},\ \{14, 23, 34\}.$$

One can check that $\tilde{\sigma}_{T_i}$ intersects $\tilde{\sigma}_{T_1} \cup \cdots \cup \tilde{\sigma}_{T_{i-1}}$ in a union of $j-1$ $(d-1)$-dimensional faces of $\tilde{\sigma}_{T_i}$, where j is the number of vertices of T_i that are less than all their neighbors. It is not hard to see that the number of noncrossing alternating trees on $[d+1]$ for which exactly j vertices are less that all their neighbors is just the Narayana number $N(d,j)$ of Exercise 6.36. It follows from the techniques of R. Stanley, *Annals of Discrete Math.* **6** (1980), 333–342 (especially Thm. 1.6), that

$$(1-x)^{d+1} \sum_{n \ge 0} i(\mathcal{P}, n) x^n = \sum_{j=1}^{d} N(d,j) x^{j-1}.$$

6.32. The Tamari lattice was first considered by D. Tamari, *Nieuw Arch. Wisk.* **10** (1962), 131–146, who proved it to be a lattice. A simpler proof of this result was given by S. Huang and D. Tamari, *J. Combinatorial Theory (A)* **13** (1972), 7–13. The proof sketched here follows J. M. Pallo, *Computer J.* **29** (1986), 171–175. For further properties of Tamari lattices and their generalizations, see P. H. Edelman and V. Reiner, *Mathematika* **43** (1996), 127–154; A. Björner and M. L. Wachs, *Trans. Amer. Math. Soc.* **349** (1997), 3945–3975 (§9); and the references given there.

6.33. **a.** Since $\mathcal{S}$ is a simplicial complex, A_n is a simplicial semilattice. It is easy to see that it is graded of rank $n-2$, the rank of an element being its cardinality (number of diagonals). To check the Eulerian property, it remains to show that $\mu(x, \hat{1}) = (-1)^{\ell(x,\hat{1})}$ for all $x \in A_n$. If $x \in A_n$ and $x \neq \hat{1}$, then x divides the polygon C into regions $C_1, \ldots, C_j$, where each C_i is a convex n_i-gon for some n_i. Let $\bar{A}_n = A_n - \{\hat{1}\}$. It follows that the interval $[x, \hat{1}]$ is isomorphic to the product $\bar{A}_{n_1} \times \cdots \times \bar{A}_{n_j}$, with a $\hat{1}$ adjoined. It follows from Exercise 5.61 (dualized) that it suffices to show that $\mu(\hat{0}, \hat{1}) = (-1)^{n-2}$. Equivalently (since we have shown that every proper interval is Eulerian), we need to show that A_n has as many elements of even rank as of odd rank. One way to proceed is as follows. For any subset B of A_n, let $\eta(B)$ denote the number of elements of B of even rank minus the number of odd rank. Label the vertices of C as $1, 2, \ldots, n$ in cyclic order. For $3 \leq i \leq n-1$, let $\mathcal{S}^*$ be the set of all elements of $\mathcal{S}$ for which either there is a diagonal from vertex 1 to some other vertex, or else such a diagonal can be adjoined without introducing any interior crossings. Given $S \in \mathcal{S}^*$, let i be the least vertex that is either connected to 1 by a diagonal or for which we can connect it to vertex 1 by a diagonal without introducing any interior crossing. We can pair S with the set S' obtained by deleting or adjoining the diagonal from 1 to i. This pairing (or involution) shows that $\eta(\mathcal{S}^*) = 0$. But $A_n - \mathcal{S}^*$ is just the interval $[T, \hat{1}]$, where T contains the single diagonal connecting 2 and n. By induction (as mentioned above) we have $\eta([T, \hat{1}]) = 0$, so in fact $\eta(A_n) = 0$.

b. See C. W. Lee, *Europ. J. Combinatorics* **10** (1989), 551–560. An independent proof was given by M. Haiman (unpublished). This polytope is called the *associahedron*. For a far-reaching generalization, see [**28**, Ch. 7] and the survey article C. W. Lee, in *DIMACS Series in Discrete Math. and Theo. Comput. Sci.* **4**, 1991, pp. 443–456.

c. Write

$$\begin{aligned} F(x, y) &= x + \sum_{n\geq 2} \sum_{i=1}^{n-1} W_{i-1}(n+1) x^n y^i \\ &= x + x^2 y + x^3(y + 2y^2) + x^4(y + 5y^2 + 5y^3) + \cdots. \end{aligned}$$

By removing a fixed exterior edge from a dissected polygon and considering the edge-disjoint union of polygons thus formed, we get the functional equation

$$F = x + y\frac{F^2}{1-F}.$$

(Compare equation (6.15).) Hence by Exercise 5.59 we have

$$\begin{aligned} F &= \sum_{m\geq 1} \frac{1}{m}[t^{m-1}] \left(x + y\frac{t^2}{1-t}\right)^m \\ &= \sum_{m\geq 1} \frac{1}{m}[t^{m-1}] \sum_{n=0}^{m} \binom{m}{n} x^n \left(y\frac{t^2}{1-t}\right)^{m-n}. \end{aligned}$$

From here it is a simple matter to obtain

$$F = x + \sum_{n \geq 2} \sum_{i=1}^{n-1} \frac{1}{n+i} \binom{n+i}{i} \binom{n-2}{i-1} x^n y^i,$$

whence

$$W_i(n) = \frac{1}{n+i} \binom{n+i}{i+1} \binom{n-3}{i}. \tag{6.74}$$

This formula goes back to T. P. Kirkman, *Phil. Trans. Royal Soc. London* **147** (1857), 217–272; E. Prouhet, *Nouvelles Annales Math.* **5** (1866), 384; and A. Cayley, *Proc. London Math. Soc. (1)* **22** (1890–1891), 237–262, who gave the first complete proof. For Cayley's proof see also [**28**, §7.3]. For a completely different proof, see Exercise 7.17. Another proof appears in D. Beckwith, *Amer. Math. Monthly* **105** (1998), 256–257.

* **d.** We have

$$h_i = \frac{1}{n-1} \binom{n-3}{i} \binom{n-1}{i+1}.$$

This result appears in C. W. Lee, *Europ. J. Combinatorics* **10** (1989), 551–560 (Thm. 6.3).

6.34. a.–d. See J. Fürlinger and J. Hofbauer, *J. Combinatorial Theory (A)* **40** (1985), 248–264, and the references given there. For (b) see also G. E. Andrews, *J. Statist. Plann. Inf.* **34** (1993), 19–22 (Cor. 1). The continued fraction (6.59) is of a type considered by Ramanujan. It is easy to show (see for instance [**1.1**, §7.1]) that

$$F(x) = \frac{\sum_{n \geq 0} (-1)^n q^{n^2} \frac{x^n}{(1-q)\cdots(1-q^n)}}{\sum_{n \geq 0} (-1)^n q^{n(n-1)} \frac{x^n}{(1-q)\cdots(1-q^n)}}.$$

e. See R. Stanley, *Ann. New York Acad. Sci.* **574** (1989), 500–535 (Example 4, p. 523). This result is closely related to Exercise 6.25(c).

6.35. These results are due to V. Welker, *J. Combinatorial Theory (B)* **63** (1995), 222–244 (§4).

6.36. a. There are $\binom{n}{k-1}\binom{n-1}{k-1}$ pairs of compositions $A : a_1 + \cdots + a_k = n + 1$ and $B : b_1 + \cdots + b_k = n$ of $n + 1$ and n into k parts. Construct from these compositions a circular sequence $w = w(A, B)$ consisting of a_1 1's, then b_1 -1's, then a_2 1's, then b_2 -1's, etc. Because n and $n + 1$ are relatively prime, this circular sequence w could have arisen from exactly k pairs (A_i, B_i) of compositions of $n + 1$ and n into k parts, viz., $A_i : a_i + a_{i+1} + \cdots + a_k + a_1 + \cdots + a_{i-1} = n + 1$ and $B_i : b_i + b_{i+1} + \cdots + b_k + b_1 + \cdots + b_{i-1} = n$, $1 \leq i \leq k$. By the second proof of Theorem 5.3.10 (or more specifically, the paragraph following it), there is exactly one way to break w into a linear sequence $\overline{w}$ such that $\overline{w}$ begins with a 1, and when this initial 1 is removed every partial sum is nonnegative. Clearly there are exactly k 1's in $\overline{w}$ (with or without its initial 1 removed) followed by a -1. This sets up a bijection between the set of all "circular

equivalence classes" $\{(A_1, B_1), \ldots, (A_k, B_k)\}$ and X_{nk}. Hence

$$X_{nk} = \frac{1}{k}\binom{n}{k-1}\binom{n-1}{k-1} = \frac{1}{n}\binom{n}{k-1}\binom{n}{k}.$$

b. For $k \geq 1$, let $X_k = X_{1k} \cup X_{2k} \cup \cdots$. Every $w \in X_k$ can be written uniquely in one of the forms (i) $1\ u\ -1\ v$, where $u \in X_j$ and $v \in X_{k-j}$ for some $1 \leq j \leq k-1$, (ii) $1\ -1\ u$, where $u \in X_{k-1}$, (iii) $1\ u\ -1$, where $u \in X_k$, and (iv) $1\ -1$ (when $k = 1$). Regarding X_k as a language as in Example 6.6.6, and replacing for notational comprehensibility 1 by α and -1 by β, conditions (i)–(iv) are equivalent to the equation

$$X_k = \sum_{j=1}^{k-1} \alpha X_j \beta X_{k-j} + \alpha\beta X_{k-1} + \alpha X_k \beta + \delta_{1k}\alpha\beta.$$

Thus if $y_k = \sum_{n \geq 1} N(n, k)x^n$, it follows that (setting $y_0 = 0$)

$$y_k = x \sum_{j=0}^{k} y_j y_{k-j} + x y_{k-1} + x y_k + \delta_{1k} x.$$

Since $F(x, t) = \sum_{k \geq 1} y_k t^k$, we get (6.60).

Narayana numbers were introduced by T. V. Narayana, *C. R. Acad. Sci. Paris* **240** (1955), 1188–1189, and considered further by him in *Sankhyā* **21** (1959), 91–98, and *Lattice Path Combinatorics with Statistical Applications*, Mathematical Expositions **23**, University of Toronto Press, Toronto, 1979 (§V.2). Further references include G. Kreweras and P. Moszkowski, *J. Statist. Plann. Inference* **14** (1986), 63–67; G. Kreweras and Y. Poupard, *Europ. J. Combinatorics* **7** (1986), 141–149; R. A. Sulanke, *Bull. Inst. Combin. Anal.* **7** (1993), 60–66; and R. A. Sulanke, *J. Statist. Plann. Inference* **34** (1993), 291–303.

6.37. Equivalent to Exercise 1.37(c). See also the nice survey R. Donaghey and L. W. Shapiro, *J. Combinatorial Theory (A)* **23** (1977), 291–301.

6.38. All these results except (f), (k), (l), and (m) appear in Donaghey and Shapiro, *loc. cit.* Donaghey and Shapiro give several additional interpretations of Motzkin numbers and state that they have found a total of about 40 interpretations. For (f), see M. S. Jiang, in *Combinatorics and Graph Theory (Hefei, 1992)*, World Scientific Publishing, River Edge, New Jersey, 1993, pp. 31–39. For (k), see A. Kuznetsov, I. Pak, and A. Postnikov, *J. Combinatorial Theory (A)* **76** (1996), 145–147. For (l), see M. Aigner, *Europ. J. Combinatorics* **19** (1998), 663–675. Aigner calls the partitions of (l) *strongly noncrossing*. Finally, for (m) see M. Klazar, *Europ. J. Combinatorics* **17** (1996), 53–68 (pp. 55–56) (and compare Exercise 1.29). Klazar's paper contains a number of further enumeration problems related to the present one that lead to algebraic generating functions; see Exercise 6.19(ss), (tt) and Exercise 6.39(o) for three of them. The name "Motzkin number" arose from the paper T. Motzkin, *Bull. Amer. Math. Soc.* **54** (1948), 352–360.

6.39. **a.** This was the definition of Schröder numbers given in the discussion of Schröder's second problem in Section 6.2.

b,e,f,h,i. These follow from (a) using the bijections of Proposition 6.2.1.

c. See D. Gouyou-Beauchamps and D. Vanquelin, *RAIRO Inform. Théor. Appl.* **22** (1988), 361–388. This paper gives some other tree representations of Schröder numbers, as well as connections with Motzkin numbers and numerous references.

d. An easy consequence of the paper of Shapiro and Stephens cited below.

g. Due to R. A. Sulanke, *J. Difference Equations and Applications*, to appear. The objects counted by this exercise are called *zebras*. See also E. Pergola and R. A. Sulanke, *J. Integer Sequences* (electronic) **1** (1998), Article 98.1.7, available at http://www.research.att.com/~njas/sequences/JIS.

j,k. See L. W. Shapiro and A. B. Stephens, *SIAM J. Discrete Math.* **4** (1991), 275–280. For (j), see also Exercise 6.17(b).

l. L. W. Shapiro and S. Getu (unpublished) conjectured that the set $\mathfrak{S}_n(2413, 3142)$ and the set counted by (k) are identical (identifying a permutation matrix with the corresponding permutation). It was proved by J. West, *Discrete Math.* **146** (1995), 247–262, that $\#\mathfrak{S}_n(2413, 3142) = r_{n-1}$. Since it is easy to see that permutations counted by (k) are 2413-avoiding and 3142-avoiding, the conjecture of Shapiro and Getu follows from the fact that both sets have cardinality r_{n-1}. Presumably there is some direct proof that the set counted by (k) is identical to $\mathfrak{S}_n(2413, 3142)$.

West also showed in Theorem 5.2 of the above-mentioned paper that the sets $\mathfrak{S}_n(1342, 1324)$ and (m) are identical. The enumeration of $\mathfrak{S}_n(1342, 1432)$ was accomplished by S. Gire, Ph.D. thesis, Université Bordeaux, 1991. The remaining seven cases were enumerated by D. Kremer, Permutations with forbidden subsequences and a generalized Schröder number, preprint. Kremer also gives proofs of the three previously known cases. She proves all ten cases using the method of "generating trees" introduced by F. R. K. Chung, R. L. Graham, V. E. Hoggatt, Jr., and M. Kleiman, *J. Combinatorial Theory (A)* **24** (1978), 382–394, and further developed by J. West, *Discrete Math.* **146** (1995), 247–262, and **157** (1996), 363–374. It has been verified by computer that there are no other pairs $(u, v) \in \mathfrak{S}_4 \times \mathfrak{S}_4$ for which $\#\mathfrak{S}_n(u, v) = r_{n-1}$ for all n.

m. This is a result of Knuth [**5.41**, Exercises 2.2.1.10–2.2.1.11, pp. 239 and 533–534]; these permutations are now called *deque-sortable*. A combinatorial proof appears in D. G. Rogers and L. W. Shapiro, in Lecture Notes in Math. **884**, Springer-Verlag, Berlin, 1981, pp. 293–303. Some additional combinatorial interpretations of Schröder numbers and many additional references appear in the preceding reference. For q-analogues of Schröder numbers, see J. Bonin, L. W. Shapiro, and R. Simion, *J. Statist. Plann. Inference* **34** (1993), 35–55.

n. See D. G. Rogers and L. W. Shapiro, Lecture Notes in Mathematics **686**, Springer-Verlag, Berlin, 1978, pp. 267–276 (§5) for simple bijections with (a) and other "Schröder structures."

o. This result is due to R. C. Mullin and R. G. Stanton, *Pacific J. Math.* **40** (1972), 167–172 (§3), using the language of "Davenport–Schinzel sequences." It is also given by M. Klazar, *Europ. J. Combinatorics* **17** (1996), 53–68 (p. 55).

p. Remove a "root edge" from the polygon of (h) and "straighten out" to obtain a noncrossing graph of the type being counted.

q,r. These results (which despite their similarity are not trivially equivalent) appear in D. G. Rogers, Lecture Notes in Math. **622**, Springer-Verlag, Berlin, 1977, pp. 175–196 (equations (38) and (39)), and are further developed in D. G. Rogers, *Quart. J. Math. (Oxford) (2)* **31** (1980), 491–506. In particular, a bijective proof that (q) and (r) are equinumerous appears in §3 of this latter reference. It is also easy to see that (p) and (r) are virtually identical problems. A further reference is D. G. Rogers and L. W. Shapiro, Lecture Notes in Mathematics **686**, Springer-Verlag, Berlin, 1978, pp. 267–276.

s. See M. Ciucu, *J. Algebraic Combinatorics* **5** (1996), 87–103, Thm. 4.1.

6.40. Note that this exercise is the "opposite" of Exercise 6.39(k), i.e., here we are counting the permutation matrices P for which not even a single new 1 can be added (using the rules of Exercise 6.39(k)). The present exercise was solved by Shapiro and Stephens in §3 of the paper cited in the solution to Exercise 6.39(k). For a less elegant form of the answer and further references, see M. Abramson and M. O. J. Moser, *Ann. Math. Statist.* **38** (1967), 1245–1254.

6.41. This result was conjectured by J. West, Ph.D. thesis, M.I.T., 1990 (Conjecture 4.2.19), and first proved by D. Zeilberger, *Discrete Math.* **102** (1992), 85–93. For further proofs and related results, see M. Bóna, 2-stack sortable permutations with a given number of runs, MSRI Preprint 1997-055; M. Bousquet-Mélou, *Electron. J. Combinatorics* **5** (1998), R21, 12 pp.; S. Dulucq, S. Gire, and J. West, *Discrete Math.* **153** (1996), 85–103; I. P. Goulden and J. West, *J. Combinatorial Theory (A)* **75** (1996), 220–242; and J. West, *Theoret. Comput. Sci.* **117** (1993), 303–313.

6.42. It's easy to see that $f(n)$ is the number of lattice paths with n steps $(1, 0)$, $(0, 1)$, and $(1, 1)$ that begin at $(0, 0)$ and end on the line $y = x$. Hence by equation (6.29) we have $F(x) = 1/\sqrt{1 - 2x - 3x^2}$. The linear recurrence is then given by Example 6.4.8(b)(ii).

It's also easy to see that $f(n) = [t^0](t^{-1} + 1 + t)^n$. For this reason $f(n)$ is called a *middle trinomial coefficient*. Middle trinomial coefficients were first considered by L. Euler, *Opuscula Analytica*, vol. 1, Petropolis, 1783, pp. 48–62, who obtained the generating function $1/\sqrt{1 - 2x - 3x^2}$. See also Problem III.217 of [**5.53**, vol. I, pp. 147 and 349]. The interpretation of $f(n)$ in terms of chess is due to K. Fabel, Problem 1413, *Feenschach* **13** (October 1974), Heft 25, p. 382; solution, **13** (May–June–July, 1975), Heft 28, p. 91. Fabel gives the solution as

$$f(n) = \sum_{i=0}^{\lfloor n/2 \rfloor} \frac{n!}{i!^2(n-2i)!},$$

but does not consider the generating function $F(x)$.

6.43. a. Define a secondary structure to be *irreducible* if either $n = 1$ or there is an edge from 1 to n. Let $t(n)$ be the number of irreducible secondary structures with n vertices, and set

$$T(x) = \sum_{n \geq 1} t(n)x^n = x + x^3 + x^4 + 2x^5 + 4x^6 + \cdots.$$

It is easy to see that $S = 1/(1-T)$ and $T = x^2S + x - x^2$. Eliminating T and solving for S yields the desired formula.

This result is due to P. R. Stein and M. S. Waterman, *Discrete Math.* **26** (1978), 261–272 (the case $m = 1$ of (10)). For further information and references, including connections with biological molecules such as RNA, see W. R. Schmitt and M. S. Waterman, *Discrete Applied Math.* **51** (1994), 317–323.

b. See A. Nkwanta, in *DIMACS Series in Discrete Math. and Theor. Comput. Sci.* **34**, 1997, pp. 137–147.

6.44. This result is due to P. H. Edelman and V. Reiner, *Graphs and Combinatorics* **13** (1997), 231–243 (Thm. 6.1).

6.45. See R. Stanley, solution to 6342, *American Math. Monthly* **90** (1983), 61–62. A labeled version of this result was given by S. Goodall, The number of labeled posets of width two, Mathematics Preprint Series LSE-MPS-46, London School of Economics, March 1993.

6.46. These remarkable results are due to G. Viennot and D. Gouyou-Beauchamps, *Advances in Appl. Math.* **9** (1988), 334–357. The subsets being enumerated are called *directed animals*. For a survey of related work, see G. Viennot, *Astérisque* **121–122** (1985), 225–246. See also the two other papers cited in the solution to Exercise 6.19(kkk), as well as the paper M. Bousquet-Mélou, *Discrete Math.* **180** (1998), 73–106. Let us also mention that Viennot and Goyou-Beauchamps show that $f(n)$ is the number of sequences of length $n-1$ over the alphabet $\{-1, 0, 1\}$ with nonnegative partial sums. Moreover, M. Klazar, *Europ. J. Combinatorics* **17** (1996), 53–68 (p. 64) shows that $f(n)$ is the number of partitions of $[n+1]$ such that no block contains two consecutive integers, and such that if $a < b < c < d$, a and d belong to the same block B_1, and b and c belong to the same block B_2, then $B_1 = B_2$.

6.47. **a.** There is a third condition equivalent to (i) and (ii) of (a) that motivated this work. Every permutation $w \in \mathfrak{S}_n$ indexes a closed Schubert cell $\bar{\Omega}_w$ in the complete flag variety $\mathrm{GL}(n, \mathbb{C})/B$. Then w is smooth if and only if the variety $\bar{\Omega}_w$ is smooth. The equivalence of this result to (i) and (ii) is implicit in K. M. Ryan, *Math. Ann.* **276** (1987), 205–244, and is based on earlier work of Lakshmibai, Seshadri, and Deodhar. An explicit statement that the smoothness of $\bar{\Omega}_w$ is equivalent to (ii) appears in V. Lakshmibai and B. Sandhya, *Proc. Indian Acad. Sci. (Math. Sci.)* **100** (1990), 45–52.

b. This generating function is due to M. Haiman (unpublished).

NOTE. It was shown by M. Bóna, *Electron. J. Combinatorics* **5**, R31 (1998), 12 pp., that there are four other inequivalent (in the sense of Exercise 6.39(l)) pairs $(u, v) \in \mathfrak{S}_4 \times \mathfrak{S}_4$ such that the number of permutations in $\mathfrak{S}_n$ that avoid them is equal to $f(n)$, viz., (1324, 2413), (1342, 2314), (1342, 2431), and (1342, 3241). (The case (1342, 2431) is implicit in Z. Stankova, *Discrete Math.* **132** (1994), 291–316.)

6.48. This result is due to M. Bóna, *J. Combinatorial Theory (A)* **80** (1997), 257–272.

6.49. **a.** Given a domino tiling of B_n, we will define a path P from the center of the left-hand edge of the middle row to the center of the right-hand edge of the middle row. Namely, each step of the path is from the center of a domino

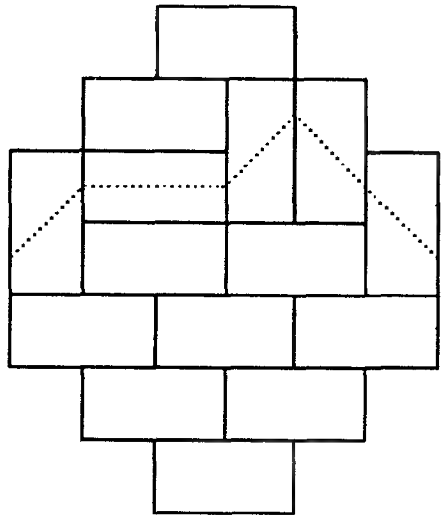


Figure 6-14. A path on the augmented Aztec diamond B_3.

edge (where we regard a domino as having six edges of unit length) to the center of another edge of the same domino D, such that the step is symmetric with respect to the center of D. One can check that for each tiling there is a unique such path P. Replace a horizontal step of P by $(1, 1)$, a northeast step by $(1, 0)$, and a southeast step by $(0, 1)$ (no other steps are possible), and we obtain a lattice path from $(0, 0)$ to (n, n) with steps $(1, 0)$, $(0, 1)$, and $(1, 1)$, and conversely any such lattice path corresponds to a unique domino tiling of B_n. This establishes the desired bijection. For instance, Figure 6-14 shows a tiling of B_3 and the corresponding path P (as a dotted line). The steps in the lattice path from $(0, 0)$ to $(3, 3)$ are $(1, 0)$, $(1, 1)$, $(1, 0)$, $(0, 1)$, $(0, 1)$.

The board B_n is called the *augmented Aztec diamond*, and its number of domino tilings was computed by H. Sachs and H. Zernitz, *Discrete Appl. Math.* **51** (1994), 171–179. The proof sketched above is based on an explication of the proof of Sachs and Zernitz due to Dana Randall (unpublished).

b. The board is called an *Aztec diamond*, and the number of tilings is now $2^{\binom{n+1}{2}}$. (Note how much larger this number is than the solution $f(n)$ to (a).) Four proofs of this result appear in N. Elkies, G. Kuperberg, M. Larsen, and J. Propp, *J. Alg. Combinatorics* **1** (1992), 111–132, 219–234. For some related work, see M. Ciucu, *J. Alg. Combinatorics* **5** (1996), 87–103, and H. Cohn, N. Elkies, and J. Propp, *Duke Math. J.* **85** (1996), 117–166. Domino tilings of the Aztec diamond and augmented Aztec diamond had actually been considered earlier by physicists, beginning with I. Carlsen, D. Grensing, and H.-Chr. Zapp, *Philos. Mag. A* **41** (1980), 777–781.

6.50. This result appears in F. R. K. Chung, R. L. Graham, J. Morrison, and A. M. Odlyzko, *Amer. Math. Monthly* **102** (1995), 113–123 (eqn. (11)). This paper contains a number of other interesting results related to pebbling.

6.51. This amazing result is due to A. Edelman, E. Kostlan, and M. Shub, *J. Amer. Math. Soc.* **7** (1994), 247–267 (Thm. 5.1).

6.52. Setting $b_k = 0$ for $k \notin \mathbb{P}$, it is easy to obtain the recurrence

$$b_n = \sum_{\substack{i+j=n \\ i<j}} b_i b_j + \left(\!\!\binom{b_{n/2}}{2}\!\!\right), \qquad n \geq 2,$$

from which (6.61) follows easily. This problem was considered by J. H. M. Wedderburn, *Ann. Math.* **24** (1922), 121–140, and I. M. H. Etherington, *Math. Gaz.* **21** (1937), 36–39, and is known as the *Wedderburn–Etherington commutative bracketing problem*. For further information and references, see [**2.3**, pp. 54–55] and H. W. Becker, *Amer. Math. Monthly* **56** (1949), 697–699.

6.53. We have

$$f(n+2) - f(n+1) = (n+2)! = (n+2)(n+1)! = (n+2)[f(n+1) - f(n)].$$

It follows that we can take $P(x) = x + 3$ and $Q(x) = x - 2$. This problem appeared as Problem B-1 in the Forty-Fifth (1984) William Lowell Putnam Competition. Of course the existence of *some* linear recurrence with polynomial coefficients satisfied by $f(n)$ follows from Theorem 6.4.9, since $n!$ is obviously P-recursive and $\sum_{n\geq 1} f(n)x^n = (1-x)^{-1} \sum_{n\geq 1} n!x^n$.

6.54. **a.** Let $y = \sum_{n\geq 0}(x^n/n!^r)$. Clearly $1/n!^r$ is P-recursive, so y is D-finite. Hence by Theorem 6.4.9, y^d is D-finite. But $y^d = \sum_{n\geq 0} S_n^{(r,d)}(x^n/n!^r)$, so $S_n^{(r,d)}/n!^r$ is P-recursive. Thus by Theorem 6.4.12 (or by a simple direct argument), $S_n^{(r,d)}$ is P-recursive. This argument appears in [**70**, Exam. 2.4] in the case $d = 2$.

b. The cases $r = 1$ and $r = 2$ are immediate from $S_n^{(1)} = 2^n$ and $S_n^{(2)} = \binom{2n}{n}$. The cases $r = 3$ and $r = 4$ are due to J. Franel, *L'Intermédiaire des Mathématiciens* **1** (1894), 45–47, and **2** (1895), 33–35. For $r = 5$ and $r = 6$, see M. A. Perlstadt, *J. Number Theory* **27** (1987), 304–309.

c. This result was conjectured by Franel in the 1895 reference above. An incomplete proof was given by T. W. Cusick, *J. Combinatorial Theory (A)* **52** (1989), 77–83. Cusick also gives a method for computing the recurrences that is simpler than Perlstadt's and that would allow the computation for some values of $r \geq 7$. The gap in Cusick's proof was pointed out by M. Stoll, who gives an elegant proof of his own in *Europ. J. Combinatorics* **18** (1997), 707–712. Franel made an additional conjecture about the form of the recurrences, but this is disproved by Perlstadt's computation. For analogous results concerning *alternating* sums of powers of binomial coefficients, see R. J. McIntosh, *J. Combinatorial Theory (A)* **63** (1993), 223–233.

6.55. **a.** See F. R. K. Chung, R. L. Graham, V. E. Hoggatt, Jr., and M. Kleiman, *J. Combinatorial Theory (A)* **24** (1978), 382–394. For more information on this fascinating subject, see C. L. Mallows, *J. Combinatorial Theory (A)* **27** (1979), 394–396; R. Cori, S. Dulucq, and G. Viennot, *J. Combinatorial Theory (A)* **43** (1986), 1–22; and S. Dulucq and O. Guibert, *Discrete Math.* **157** (1996), 91–106. This last paper contains some additional references. See also Exercise 6.19(mm).

b. Let $f(k) = 1/(k-1)!k!(k+1)!$. Thus

$$B(n) = \binom{n+1}{1}^{-1}\binom{n+1}{2}^{-1}(n+1)!^3 \sum_{k=1}^{n} f(k)f(n+1-k).$$

Clearly $f(k)$ is P-recursive, so by Proposition 6.4.3 and Theorem 6.4.9 the function $\sum_{k=1}^{n} f(k)f(n+1-k)$ is P-recursive. It is then easy to see (e.g., from Theorem 6.4.12, though a simple direct argument can also be given) that $B(n)$ is P-recursive. This result can also be deduced from the general theory presented in [**75**].

c. According to Chung *et al.* (reference in (a)), P. S. Bruckman derived from (6.62) that

$$\begin{aligned}(n+1)&(n+2)(n+3)(3n-2)B(n)\\ &= 2(n+1)(9n^3+3n^2-4n+4)B(n-1)\\ &\quad + (3n-1)(n-2)(15n^2-5n-14)B(n-2)\\ &\quad + 8(3n+1)(n-2)^2(n-3)B(n-3),\end{aligned}$$

for $n \geq 4$.

6.56. a. Using the notation and techniques of Chapter 7 (see Corollary 7.23.12), we have

$$A_k(n) = \sum_{\substack{\lambda \vdash n \\ \ell(\lambda)<k}} (f^\lambda)^2,$$

where f^λ denotes the number of standard Young tableaux of shape λ. Using the hook-length formula for f^λ (Corollary 7.21.6), we get an explicit formula for $A_k(n)$ from which the techniques of [**75**] (especially §3.3) imply that A_k is P-recursive. Another approach is given by I. M. Gessel, *J. Combinatorial Theory (A)* **53** (1990), 257–285 (§7).

b. These conjectures are due to F. Bergeron, L. Favreau, and D. Krob, *Discrete Math.* **139** (1995), 463–468. They were obtained using the Maple package *gfun* and some related tools developed by S. Plouffe.

* **c.** Very little is known about the function A_v for arbitrary $v \in \mathfrak{S}_k$. For instance, it is conjectured that $\lim_{n\to\infty} A_v(n)^{1/n}$ exists (and is finite), but it is not even known whether $A_v(n) < c^n$ for some constant $c > 0$ depending on v. For further work on the function A_v, see Exercises 6.19(ee,ff), 6.39(l), 6.47, and 6.48

6.57. Let $Q(x) = 1+\alpha_1 x+\cdots+\alpha_d x^d = \prod_{i=1}^{k}(1-\gamma_i x)^{d_i}$, where the γ_i's are distinct nonzero complex numbers and $d_i > 0$. We claim that the answer is k, the number of *distinct* zeros of $Q(x)$. By Theorem 4.1.1 we have $f(n) = \sum_{i=1}^{k} P_i(n)\gamma_i^n$ for some polynomials $0 \neq P_i(n) \in \mathbb{C}[n]$. Let E be the unit shift operator, given by $(Eg)(n) = g(n+1)$. Let

$$g(n) = [P_k(n)E - \gamma_k P_k(n+1)]f.$$

It's easy to see that $g(n) = \sum_{i=1}^{k-1} Q_i(n)\gamma_i^n$ for some polynomials $Q_i(n) \in \mathbb{C}[n]$. Hence by induction on k (the case $k=1$ being trivial), there is an operator $\Omega = H_{k-1}(n)E^{k-1}+\cdots+H_1(n)E+H_0(n)$, where $H_j(n) \in \mathbb{C}[n]$, satisfying

$\Omega g = 0$. Hence $\Omega \cdot [P_k(n)E - \gamma_k P_k(n+1)]f = 0$, a (nonzero) homogeneous linear recurrence of order at most k with polynomial coefficients satisfied by f.

It remains to show that f cannot satisfy a recurrence of order less than k. Suppose to the contrary that

$$T_{k-1}(n)f(n+k-1) + T_{k-2}f(n+k-2) + \cdots + T_0(n)f(n) = 0$$

for all $n \geq 0$, where $T_i(n) \in \mathbb{C}[n]$ and not all $T_i = 0$. Then

$$\begin{aligned} 0 &= \sum_{j=0}^{k-1} T_j(n) \sum_{i=1}^{k} P_i(n+j)\gamma_i^{n+j} \\ &= \sum_{i=1}^{k} \gamma_i^n \sum_{j=0}^{k-1} T_j(n) P_i(n+j)\gamma_i^j. \end{aligned}$$

Since the functions γ_i^n are linearly independent over $\mathbb{C}(n)$ (e.g., by Theorem 4.1.1), we have

$$\sum_{j=0}^{k-1} T_j(n)P_i(n+j)\gamma_i^j = 0, \qquad 1 \leq i \leq k.$$

Since the Vandermonde matrix $[\gamma_i^j]$ has nonzero determinant, we have $T_j(n)P_i(n+j) = 0$ for all $0 \leq j \leq k-1$, $1 \leq i \leq k$, and $n \geq 0$. Since each P_i and some T_j have only finitely many zeros, we have reached a contradiction.

6.58. Let r be the maximum nonnegative integer zero of $P_e(n)$ (where we set $r = -1$ if $P_e(n)$ has no such zero). Once we know $f(0), \ldots, f(r+e)$, then all other values of $f(n)$ are uniquely determined by (6.34). Hence $\mathcal{V}$ has the same dimension as the space of all $(r+e+1)$-tuples $(f(0), f(1), \ldots, f(r+e))$, where $f \in \mathcal{V}$. If we substitute $n = 0, 1, \ldots, r$ in (6.34), then we obtain $r+1$ homogeneous linear equations $E_0, E_1, \ldots, E_r$ in the $r+e+1$ unknowns $f(0), f(1), \ldots, f(r+e)$. If n is not a zero of $P_e(n)$, then the equation E_n involves $f(n+e)$ with a nonzero coefficient, while the equations $E_0, E_1, \ldots, E_{n-1}$ do not involve $f(n+e)$. Thus the system $(E_0, E_1, \ldots, E_n)$ has rank one more than the rank of the system $(E_0, E_1, \ldots, E_{n-1})$. It follows that the rank of the system $(E_0, E_1, \ldots, E_r)$ is at least $r+1-m$. On the other hand, the rank is at most the number $r+1$ of equations. Since the dimension of the solution space is the number $r+e+1$ of unknowns minus the rank, we get $e \leq \dim \mathcal{V} \leq e+m$, as desired. Given e and $P_e(n)$, it is easy to arrange for any value of $\dim \mathcal{V}$ in this range to occur. (In fact, since one can specify finitely many values of a polynomial arbitrarily, the system $(E_0, E_1, \ldots, E_r)$ can also be specified arbitrarily, except for being consistent with $P_e(n)$ having m zeros in $\mathbb{N}$, the largest being r.)

6.59. Let $u = \sec x$. Suppose that u satisfies the differential equation (6.31). Now $u' = u\sqrt{u^2-1}$, then $u'' = u^3 + u^2 - u$, and in general by induction it is easily seen that $u^{(2i+1)} = L_i(u)\sqrt{u^2-1}$ and $u^{(2i)} = M_i(u)$, where L_i and M_i are polynomials (with complex coefficients), both of degree $2i+1$. Making these substitutions into (6.31) yields a *nonzero* polynomial equation in x, u, and $\sqrt{u^2-1}$ satisfied by u. Hence u is algebraic, which is easily seen to be impossible (e.g., by Exercise 6.1). This argument appears in [**70**, Exam. 2.5]. An earlier proof was given by L. Carlitz, *J. Reine Angew. Math.* **214/215** (1964), 184–191

(Thm. 4). Another proof follows from the result within the theory of differential equations that an analytic D-finite series cannot have infinitely many poles. See [**70**, §4(a)] for a stronger result. Yet another way to see that $\sec x$ is not D-finite is to appeal to the difficult Exercise 6.60. An argument similar to the one given above can be applied to $\sqrt{\log(1+x^2)}$. See [**70**, Exam. 2.6] for further details.

6.60. Suppose that $u := y'/y$ is algebraic. Repeatedly differentiating the equation $y' = uy$ and using induction shows that $y^{(k)} = P_k(u, u', \dots)y$, where P_k is a polynomial (over $\mathbb{C}$) in $u, u', \dots$. Since u is algebraic, all of the series $P_k(u, u', \dots)$ lie in the field $\mathbb{C}(x, u)$ (using (6.12)) and hence satisfy some linear dependence relation $\sum_{k=0}^{m} f_k(x)P_k = 0$, where $f_k(x) \in \mathbb{C}(x)$. Thus $\sum f_k(x)y^{(k)} = 0$, so y is D-finite.

The converse is considerably more difficult. It was first proved explicitly by W. A. Harris, Jr., and Y. Sibuya, *Advances in Math.* **58** (1985), 119–132, though it actually follows from an earlier result of S. Morrison appearing in P. Blum, *Amer. J. Math.* **94** (1972), 676–684 (Thm. 3). The result of Harris and Sibuya was successively generalized by W. A. Harris, Jr., and Y. Sibuya, *Proc. Amer. Math. Soc.* **97** (1986), 207–211, and S. Sperber, *Pacific J. Math.* **124** (1986), 249–256, culminating in M. F. Singer, *Trans. Amer. Math. Soc.* **295** (1986), 753–763.

6.61. This is a result of L. Lipshitz, *J. Algebra* **113** (1988), 373–378. Earlier proofs by I. Gessel, *Utilitas Math.* **19** (1981), 247–254, and D. Zeilberger, *J. Math. Anal. Appl.* **85** (1982), 114–145 (Thm. 11), were incomplete, though Zeilberger later completed his approach. A more general result was later proved by L. Lipshitz, *J. Algebra* **122** (1989), 353–373 (Thm. 2.7), using his generalization of D-finiteness to several variables.

6.62. Suppose that y satisfies an appropriate differential equation $\mathcal{E}$ of order e. Applying to $\mathcal{E}$ an element of the Galois group G of the irreducible polynomial P of which y is a root, we see that every conjugate of y also satisfies $\mathcal{E}$. It's easy to see that the space of fractional power series (or even fractional Laurent series) solutions to $\mathcal{E}$ has complex dimension at most e. Hence $e \geq \dim V$.

It remains to show that y satisfies *some* equation $\mathcal{E}$ whose degree is equal to $\dim V$. Let L be the splitting field of P. The derivation d/dx extends uniquely to L. G leaves V invariant, and so we have a representation $G \to \mathrm{GL}(V)$. For any $\mathbb{C}$-basis $z_1, \dots, z_m$ of V, let

$$T(Y) = \frac{\mathrm{Wr}(Y, z_1, \dots, z_m)}{\mathrm{Wr}(z_1, \dots, z_m)},$$

where Y is an indeterminate and Wr denotes the Wronskian determinant. For any $\sigma \in G$, apply σ to the coefficients of $T(Y)$ and call the resulting differential polynomial $T^\sigma(Y)$. We then have

$$\begin{aligned} T^\sigma(Y) &= \frac{\mathrm{Wr}(Y, \sigma z_1, \dots, \sigma z_m)}{\mathrm{Wr}(\sigma z_1, \dots, \sigma z_m)} \\ &= \frac{\det([\sigma]) \cdot \mathrm{Wr}(Y, z_1, \dots, z_m)}{\det([\sigma]) \cdot \mathrm{Wr}(z_1, \dots, z_m)} \\ &= T(Y), \end{aligned}$$

where $[\sigma]$ is the matrix of σ with respect to the basis $z_1, \dots, z_m$. Therefore the

coefficients of $T(Y)$ are left invariant by G and so lie in $\mathbb{C}(x)$. $T(Y)$ is clearly nonzero (since $z_1, \ldots, z_m$ are linearly independent), vanishes at the roots of P, and has order equal to dim V, as desired.

The above argument is due to B. Dwork and later independently M. F. Singer (private communications).

6.63. **b.** It was shown by K. G. J. Jacobi, *J. Reine Angew. Math.* (= *Crelle's J.*) **36** (1847), 97–112, that the series $y = 1 + 2\sum_{n\geq 1} x^{n^2}$ satisfies the ADE

$$\left(y^2 z_3 - 15 y z_1 z_2 + 30 z_1^3\right)^2 + 32\left(y z_2 - 3 z_1^2\right)^3 - y^{10}\left(y z_2 - 3 z_1^2\right)^2 = 0,$$

where $z_1 = xy'$, $z_2 = xy' + x^2y''$, $z_3 = xy' + 3x^2y'' + x^3y'''$.

c. The strongest related result known to date is due to L. Lipshitz and L. A. Rubel, *Amer. J. Math.* **108** (1986), 1193–1214 (Thm. 4.1). Their result shows in particular that y cannot have Hadamard gaps. In other words, if y satisfies an ADE, then there does not exist $r > 1$ such that $n_{i+1}/n_i > r$ for all i. It is open whether some series of the form $\sum_n b_n x^{n^3}$, with each $b_n \neq 0$, can satisfy an ADE. In fact, it does not seem to be known whether the series $\sum_n x^{n^3}$ satisfies an ADE. (It seems likely that $\sum_n b_n x^{n^3}$ does not satisfy an ADE, since otherwise there would be a completely unexpected result about representing integers as sums of cubes.)

6.64. **a.** The series represented by both sides is just the sum of all words $w \in \{x, y\}^*$ whose letters alternate (i.e., no two consecutive x's or y's).

b. Let $u = (1 - xy)^{-1}$ and $v = (1 - yx)^{-1}$. Note that $(1 - yx)y = y(1 - xy)$, and therefore $yu = vy$. Thus

$$\begin{aligned}(1 + x)v(1 + y) &= v + xv + yu + xyu\\ &= v + xv + yu + u - (1 - xy)u\\ &= u + v + xv + yu - 1.\end{aligned}$$

This last expression is symmetric with respect to the permutation (written in disjoint cycle form) $(x, y)(u, v)$, so an affirmative answer follows. This argument is due to S. Fomin (private communication), and the problem itself was motivated by the paper S. Fomin, *J. Combinatorial Theory (A)* **72** (1995), 277–292 (proof of Thm. 1.2).

c. An affirmative answer is due to D. Krob, in *Topics in Invariant Theory* (M.-P. Malliavin, ed.), Lecture Notes in Math. **1478**, Springer-Verlag, Berlin/Heidelberg/New York, 1991, pp. 215–243. A short discussion also appears in [55, §8].

6.65. **a.** **First Solution.** Suppose that $S = \sum_{n\geq 1} x^n y^n$ were rational. By Theorem 6.5.7, there exist $N \times N$ matrices A and B (for some N) such that $(A^i B^j)_{1N} = \delta_{ij}$ for all $i, j \geq 1$. Form the *commutative* generating function

$$\begin{aligned}F(s, t) &= \sum_{i,j\geq 1} (A^i B^j)_{1N} s^i t^j\\ &= [As(I - As)^{-1} Bt(I - Bt)^{-1}]_{1N}.\end{aligned}$$

By an argument as in Exercise 4.8(a), we see that $F(s, t)$ is a rational function

of s and t with denominator $\det(I - As) \cdot \det(I - At)$. On the other hand,

$$F(s,t) = \sum_{i,j\geq 1} \delta_{ij} s^i t^j = \frac{st}{1-st},$$

so that $1 - st$ is a factor of the denominator of $F(s,t)$, a contradiction (e.g., by the fact that $K[s,t]$ is a unique factorization domain).

Second Solution. Let A, B, and N be as above. By the Cayley–Hamilton theorem, we have $A^N = a_{N-1}A^{N-1} + \cdots + a_1 A + a_0 I$ for certain constants a_i. Hence

$$1 = (A^N B^N)_{1N} = \sum_{i=0}^{N-1} a_i (A^i B^N)_{1N} = 0,$$

a contradiction. This elegant argument is due to P. Hersh.

b. This is an example of an "iteration lemma" or "pumping lemma" for rational series, and is due to G. Jacob, *J. Algebra* **63** (1980), 389–412. For some stronger results and additional references, see C. Reutenauer, *Acta Inf.* **13** (1980), 189–197. See also [**5**, Ch. III, Thm. 4.1][**43**, Thm. 8.23].

c. This pumping lemma is a result of J. Jaffe, *SIGACT News* (1978), 48–49. See also [**62**, Thm. 3.14]. The paper of Jaffe mentions some further characterizations of rational languages, the earliest due to A. Nerode.

6.66. By Theorem 6.5.7, there are homomorphisms $\mu : X^+ \to K^{m\times m}$ and $\nu : X^+ \to K^{n\times n}$ such that $\langle S, w\rangle = \mu(w)_{1m}$ and $\langle T, w\rangle = \nu(w)_{1n}$ for all $w \in X^+$. Define $\mu \otimes \nu : X^+ \to K^{mn\times mn}$ by $(\mu \otimes \nu)(w) = \mu(w) \otimes \nu(w)$, where the latter $\otimes$ denotes tensor (Kronecker) product of matrices. Then $\mu \otimes \nu$ is a homomorphism of monoids satisfying $(\mu\otimes\nu)(w)_{1,mn} = \langle S, w\rangle \cdot \langle T, w\rangle$ for all $w \in X^+$, and the proof follows from Theorem 6.5.7. This result is due to M. P. Schützenberger [**65**, Property 2.2]. See also for instance [**63**, Thm. II.4.4]. (Schützenberger assumes only that K is a semiring, in which case the proof is considerably more difficult.)

6.67. Let $x = x_0 + x_1 + \cdots + x_{b-1}$ and $y = x_1 + 2x_2 + \cdots + (b-1)x_{b-1}$. Then one sees easily that

$$S = bSx + (x - x_0)(1-x)^{-1}y + y,$$

whence

$$S = [(x - x_0)(1-x)^{-1}y + y](1 - bx)^{-1}.$$

This result is a slight variation of [**5**, Exer. 4.4, p. 19].

6.68. A proof of this result is given in M. W. Davis and M. D. Shapiro, Coxeter groups are automatic, Ohio State University, 1991, preprint. However, the proof of a result called the parallel-wall property is incomplete. Subsequently B. Brink and R. B. Howlett, *Math. Ann.* **296** (1993), 179–190, asserted that their Thm. 2.8 implies the parallel-wall property, thereby completing the proof of Davis and Shapiro. Other proofs were given by H. Eriksson, Ph.D. thesis, Kungl. Tekniska Högskolan, 1994 (Thm. 73) and P. Headley, Ph.D. thesis, University of Michigan, 1994 (Thm. VII.12).

6.69. For this result and a number of related results, see A. Björner and C. Reutenauer, *Theoret. Comput. Sci.* **98** (1992), 53–63.

6.70. This is something of a trick question – clearly (0, 0) is a solution.

6.71. Let $\alpha, \beta \in K$. Let $s_1 = u$ be a component of a proper algebraic system $\mathcal{S}$ in variables $s_1, s_2, \ldots, s_j$, and let $t_1 = v$ be a component of a proper algebraic system $\mathcal{T}$ in variables $t_1, t_2, \ldots, t_k$. Then $z = \alpha u + \beta v$ is a component of the proper algebraic system consisting of $\mathcal{S}$, $\mathcal{T}$, and the equation $z = \alpha s_1 + \beta t_1$. Hence $\alpha u + \beta v$ is algebraic. A similar argument works for uv.

It remains to show that if u is algebraic and u^{-1} exists (i.e., $\langle u, 1\rangle \neq 0$), then u^{-1} is algebraic. Suppose $\langle u, 1\rangle = \alpha \neq 0$. Let $v = u^{-1}$, $u' = u - \alpha$, and $v' = v - \alpha^{-1}$, so $\langle u', 1\rangle = \langle v', 1\rangle = 0$. Then $v' = -\alpha^{-2}u' - \alpha^{-1}v'u'$, from which it is immediate that v' (and therefore v) is algebraic whenever u' (or u) is algebraic.

For additional "closure properties" of algebraic series, see [**63**, Ch. IV.3] and [**24**].

6.72. **a.** This pumping lemma for algebraic languages is due to Y. Bar-Hillel, *Z. Phonetik, Sprachwiss. u. Kommunicationsforschung* **14** (1961), 143–172 (Thm. 4.1); reprinted in Y. Bar-Hillel, *Language and Information*, Addison-Wesley, Reading, Massachusetts, 1964, p. 130. See also [**62**, Thm. 3.13]. For additional methods of showing that some languages are not algebraic, see [**63**, Exers. IV.2.4 and IV.5.8].

b. Easy.

6.73. **a.** Let D be the Dyck language of Example 6.6.6. Let $\bar{D}$ be the set of all words obtained from words in D by interchanging x's and y's. A simple combinatorial argument shows that $S = (1 - D^+ - \bar{D}^+)^{-1}$. Since D and $\bar{D}$ are algebraic, it follows immediately that S is algebraic. More explicitly, if $S' = S - 1$ then we have (using (6.51))

$$\begin{aligned} S' &= (S'+1)D^+ + (S'+1)\bar{D}^+ \\ D^+ &= x(D^+ + 1)y(D^+ + 1) \\ \bar{D}^+ &= y(\bar{D}^+ + 1)x(\bar{D}^+ + 1). \end{aligned}$$

b. Define a series Q by $Q = 1 + xQyQ + yQxQ$. Then Q is algebraic, and the *support* of Q is the language S, as noted by J. Berstel, *Transductions and Context-Free Languages*, Teubner, 1979 (Thm. II.3.7 on p. 41). However, this leaves open the question of whether $S - 1$ itself can be defined by a single equation (or even two equations).

6.74. Every word $w = w_1 w_2 \cdots w_n$ in the alphabet $\{x_1, x_1^{-1}, \ldots, x_k, x_k^{-1}\}$ that reduces to 1 can be uniquely written as a product of irreducible such words. If $g(n)$ is the number of such irreducible words of length n and if $G(t) = \sum_{n\geq 1} g(n)t^n$, then it follows that $F = 1/(1-G)$. If an irreducible word u begins with a letter $y = x_i$ or x_i^{-1}, then it must end in y^{-1} [why?]. If $u = yvy^{-1}$, then v can be any word reducing to 1 whose irreducible components don't begin with y^{-1}. The generating function for irreducible words not beginning with y^{-1} is $\frac{2k-1}{k}G(t)$, so the generating function for sequences of such words is

$$\left(1 - \frac{2k-1}{k}G(t)\right)^{-1}.$$

Since the word u is two letters longer than v and there are $2k$ choices for y, there follows

$$G(t) = \frac{2kt^2}{1 - \frac{2k-1}{k}G(t)}.$$

From this it is easy to solve for $G(t)$ and then $F(t)$.

The argument given above is due to D. Grabiner. Another elegant solution by A. J. Schwenk appears in *Amer. Math. Monthly* **92** (1985), 670–671. The problem was formulated originally by M. Haiman and D. Richman for the case $k = 2$ in *Amer. Math. Monthly* **91** (1984), 259, though Schenk notes that his solution carries over to any k. The problem of Haiman and Richman provided the motivation for the paper [**35**], from which we obtained Theorem 6.7.1 and Corollary 6.7.2.

This problem was solved independently by physicists, in the context of random walks on the Bethe lattice. See A. Giacometti, *J. Phys. A: Math. Gen.* **28** (1995), L13–L17, and the references given there.

7

Symmetric Functions

7.1 Symmetric Functions in General

The theory of symmetric functions has many applications to enumerative combinatorics, as well as to such other branches of mathematics as group theory, Lie algebras, and algebraic geometry. Our aim in this chapter is to develop the basic combinatorial properties of symmetric functions; the connections with algebra will only be hinted at in Sections 7.18 and 7.24, Appendix 2, and in some exercises.

Let $x = (x_1, x_2, \ldots)$ be a set of indeterminates, and let $n \in \mathbb{N}$. A *homogeneous symmetric function of degree* n over a commutative ring R (with identity) is a formal power series

$$f(x) = \sum_{\alpha} c_\alpha x^\alpha,$$

where (a) α ranges over all weak compositions $\alpha = (\alpha_1, \alpha_2, \ldots)$ of n (of infinite length), (b) $c_\alpha \in R$, (c) x^α stands for the monomial $x_1^{\alpha_1} x_2^{\alpha_2} \ldots$, and (d) $f(x_{w(1)}, x_{w(2)}, \ldots) = f(x_1, x_2, \ldots)$ for every permutation w of the positive integers $\mathbb{P}$. (A symmetric function of degree 0 is just an element of R.) Note that the term "symmetric function" is something of a misnomer; $f(x)$ is not regarded as a function but rather as a formal power series. Nevertheless, for historical reasons we adhere to the above terminology.

The set of all homogeneous symmetric functions of degree n over R is denoted Λ_R^n. Clearly if $f, g \in \Lambda_R^n$ and $a, b \in R$, then $af + bg \in \Lambda_R^n$; in other words, Λ_R^n is an R-module. For our purposes it will suffice to take $R = \mathbb{Q}$ (or sometimes $\mathbb{Q}$ with some indeterminates adjoined), so $\Lambda_{\mathbb{Q}}^n$ is a $\mathbb{Q}$-vector space. For the sake of convenience, then, and because some readers are doubtless more comfortable with vector spaces than with modules, we will henceforth work over $\mathbb{Q}$, though this is not the most general approach.

If $f \in \Lambda_{\mathbb{Q}}^m$ and $g \in \Lambda_{\mathbb{Q}}^n$, then it is clear that $fg \in \Lambda_{\mathbb{Q}}^{m+n}$ (where fg is a product of formal power series). Hence if we define

$$\Lambda_{\mathbb{Q}} = \Lambda_{\mathbb{Q}}^0 \oplus \Lambda_{\mathbb{Q}}^1 \oplus \cdots \quad \text{(vector space direct sum)} \tag{7.1}$$

(so the elements of $\Lambda_{\mathbb{Q}}$ are power series $f = f_0+f_1+\cdots$ where $f_n \in \Lambda^n_{\mathbb{Q}}$ and all but finitely many $f_n = 0$), then $\Lambda_{\mathbb{Q}}$ has the structure of a $\mathbb{Q}$-algebra (i.e., a ring whose operations are compatible with the vector space structure), called the *algebra* (over $\mathbb{Q}$) *of symmetric functions*. Note that the algebra $\Lambda_{\mathbb{Q}}$ is commutative and has an identity element $1 \in \Lambda^0_{\mathbb{Q}}$. The decomposition (7.1) in fact gives $\Lambda_{\mathbb{Q}}$ the structure of a *graded algebra*, meaning that if $f \in \Lambda^m_{\mathbb{Q}}$ and $g \in \Lambda^n_{\mathbb{Q}}$, then $fg \in \Lambda^{m+n}_{\mathbb{Q}}$. From now on we suppress the subscript $\mathbb{Q}$ and write simply Λ^n and Λ for $\Lambda^n_{\mathbb{Q}}$ and $\Lambda_{\mathbb{Q}}$. Note, however, that in the outside literature Λ usually denotes $\Lambda_{\mathbb{Z}}$.

A central theme in the theory of symmetric functions is to describe various bases of the vector space Λ^n and the transition matrices between pairs of these bases. We will begin with four "simple" bases. In Sections 7.10–7.19 we consider a less obvious basis which is crucial for the deeper parts of the theory.

7.2 Partitions and Their Orderings

Recall from Section 1.3 that a *partition* λ of a nonnegative integer n is a sequence $(\lambda_1, \ldots, \lambda_k) \in \mathbb{N}^k$ satisfying $\lambda_1 \geq \cdots \geq \lambda_k$ and $\sum \lambda_i = n$. Any $\lambda_i = 0$ is considered irrelevant, and we identify λ with the *infinite* sequence $(\lambda_1, \ldots, \lambda_k, 0, 0, \ldots)$. We let $\mathrm{Par}(n)$ denote the set of all partitions of n, with $\mathrm{Par}(0)$ consisting of the empty partition $\emptyset$ (or the sequence $(0, 0, \ldots)$), and we let

$$\mathrm{Par} := \bigcup_{n\geq 0} \mathrm{Par}(n).$$

For instance (writing for example 4211 as short for $(4, 2, 1, 1, 0, \ldots)$),

$$\begin{aligned}
\mathrm{Par}(1) &= \{1\}\\
\mathrm{Par}(2) &= \{2, 11\}\\
\mathrm{Par}(3) &= \{3, 21, 111\}\\
\mathrm{Par}(4) &= \{4, 31, 22, 211, 1111\}\\
\mathrm{Par}(5) &= \{5, 41, 32, 311, 221, 2111, 11111\}.
\end{aligned}$$

If $\lambda \in \mathrm{Par}(n)$, then we also write $\lambda \vdash n$ or $|\lambda| = n$. The number of parts of λ (i.e., the number of nonzero λ_i) is the *length* of λ, denoted $\ell(\lambda)$. Write $m_i = m_i(\lambda)$ for the number of parts of λ that equal i, so in the notation of Section 1.3 we have $\lambda = \langle 1^{m_1} 2^{m_2} \cdots \rangle$, which we sometimes abbreviate as $1^{m_1}2^{m_2}\cdots$. Also recall from the discussion of entry 10 of the Twelvefold Way in Section 1.4 that the *conjugate partition* $\lambda' = (\lambda'_1, \lambda'_2, \ldots)$ of λ is defined by the condition that the Young (or Ferrers) diagram of λ' is the transpose of the Young diagram of λ; equivalently, $m_i(\lambda') = \lambda_i - \lambda_{i+1}$. Note that $\lambda'_1 = \ell(\lambda)$ and $\lambda_1 = \ell(\lambda')$.

Three partial orderings on partitions play an important role in the theory of symmetric functions. We first define $\mu \subseteq \lambda$ for any $\mu, \lambda \in \mathrm{Par}$ if $\mu_i \leq \lambda_i$ for all i. If we identify a partition with its (Young) diagram, then the partial order

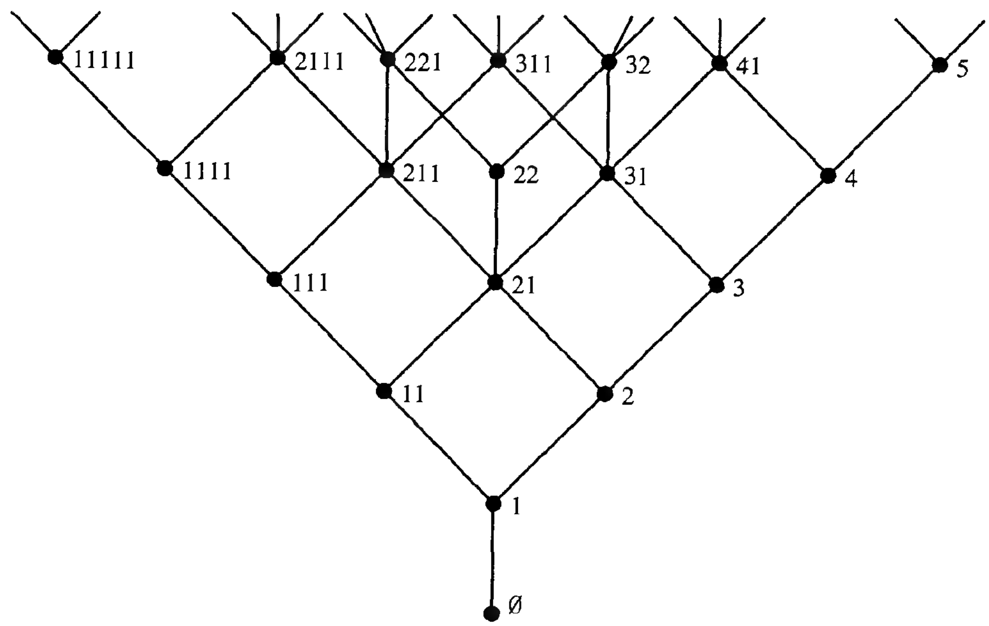


Figure 7-1. Young's lattice.

$\subseteq$ is given simply by containment of diagrams. The set Par with the partial order $\subseteq$ is called *Young's lattice* Y and (as mentioned in Exercise 3.63) is isomorphic to $J_f(\mathbb{N}^2)$. The rank of a partition λ in Young's lattice is equal to the sum $|\lambda|$ of its parts, so the rank-generating function by equation (1.30) is given by

$$F(Y, x) = \prod_{i \geq 1} (1 - x^i)^{-1}.$$

See Figure 7-1 for the first six levels of Young's lattice.

The second partial order is defined only on Par(n) for each $n \in \mathbb{N}$, and is called *dominance order* (also known as *majorization order* or the *natural order*), denoted $\leq$. Namely, if $|\mu| = |\lambda|$ then define $\mu \leq \lambda$ if

$$\mu_1 + \mu_2 + \cdots + \mu_i \leq \lambda_1 + \lambda_2 + \cdots + \lambda_i \qquad \text{for all } i \geq 1.$$

For the reader's benefit we state the following basic facts about dominance order, though we have no need for them here.

- (Par(n), $\leq$) is a lattice.
- The map $\lambda \mapsto \lambda'$ is an anti-automorphism of (Par(n), $\leq$) (so Par(n) is self-dual under dominance order).
- (Par(n), $\leq$) is a chain if and only if $n \leq 5$.
- (Par(n), $\leq$) is graded if and only if $n \leq 6$.

- For $\lambda = (\lambda_1, \dots, \lambda_n) \vdash n$, let $\mathcal{P}_\lambda$ denote the convex hull in $\mathbb{R}^n$ of all permutations of the coordinates of λ. Then $(\mathrm{Par}(n), \leq)$ is isomorphic to the set of $\mathcal{P}_\lambda$'s ordered by inclusion.

For the Möbius function of $(\mathrm{Par}(n), \leq)$ see Exercise 3.55, and for some further properties see Exercise 7.2.

For our final partial order, also on $\mathrm{Par}(n)$, it suffices to take any linear order compatible with (i.e., a linear extension of) dominance order. The most convenient is *reverse lexicographic order*, denoted $\overset{R}{\leq}$. Given $|\lambda| = |\mu|$, define $\mu \overset{R}{\leq} \lambda$ if either $\mu = \lambda$, or else for some i,

$$\mu_1 = \lambda_1, \dots, \quad \mu_i = \lambda_i, \quad \mu_{i+1} < \lambda_{i+1}.$$

For instance, the order $\overset{R}{>}$ on Par(6) is given by

$$6 \overset{R}{>} 51 \overset{R}{>} 42 \overset{R}{>} 411 \overset{R}{>} 33 \overset{R}{>} 321 \overset{R}{>} 3111 \overset{R}{>} 222 \overset{R}{>} 2211 \overset{R}{>} 21111 \overset{R}{>} 111111.$$

On the other hand, in dominance order the partitions 33 and 411, as well as 3111 and 222, are incomparable.

We would like to make one additional definition related to partitions. Define the *rank* of a partition $\lambda = (\lambda_1, \lambda_2, \dots)$, denoted $\mathrm{rank}(\lambda)$, to be the largest integer i for which $\lambda_i \geq i$. Equivalently, $\mathrm{rank}(\lambda)$ is the length of the main diagonal of the diagram of λ or the side length of the Durfee square of λ (defined in the solution to Exercise 1.23(b)). Note that $\mathrm{rank}(\lambda) = \mathrm{rank}(\lambda')$.

7.3 Monomial Symmetric Functions

Given $\lambda = (\lambda_1, \lambda_2, \dots) \vdash n$, define a symmetric function $m_\lambda(x) \in \Lambda^n$ by

$$m_\lambda = \sum_\alpha x^\alpha,$$

where the sum ranges over all *distinct* permutations $\alpha = (\alpha_1, \alpha_2, \dots)$ of the entries of the vector $\lambda = (\lambda_1, \lambda_2, \dots)$. For instance,

$$\begin{aligned} m_\emptyset &= 1 \\ m_1 &= \sum_i x_i \\ m_2 &= \sum_i x_i^2 \\ m_{11} &= \sum_{i<j} x_i x_j. \end{aligned}$$

We call m_λ a *monomial symmetric function*. Clearly if $f = \sum_\alpha c_\alpha x^\alpha \in \Lambda^n$ then $f = \sum_{\lambda \vdash n} c_\lambda m_\lambda$. It follows that the set $\{m_\lambda : \lambda \vdash n\}$ is a (vector space) basis for

Λ^n, and hence that

$$\dim \Lambda^n = p(n),$$

the number of partitions of n. Moreover, the set $\{m_\lambda : \lambda \in \text{Par}\}$ is a basis for Λ.

7.4 Elementary Symmetric Functions

Next we define the *elementary symmetric functions* e_λ for $\lambda \in$ Par by the formulas

$$\begin{aligned} e_n &= m_{1^n} = \sum_{i_1<\cdots<i_n} x_{i_1}\cdots x_{i_n}, \quad n \geq 1 \qquad (\text{with } e_0 = m_\emptyset = 1) \\ e_\lambda &= e_{\lambda_1} e_{\lambda_2} \cdots, \qquad \text{if } \lambda = (\lambda_1, \lambda_2, \ldots). \end{aligned} \tag{7.2}$$

If $A = (a_{ij})_{i,j\geq 1}$ is an integer matrix with finitely many nonzero entries and with row and column sums

$$\begin{aligned} r_i &= \sum_j a_{ij} \\ c_j &= \sum_i a_{ij}, \end{aligned}$$

then define the *row-sum vector* $\text{row}(A)$ and *column-sum vector* $\text{col}(A)$ by

$$\begin{aligned} \text{row}(A) &= (r_1, r_2, \ldots) \\ \text{col}(A) &= (c_1, c_2, \ldots). \end{aligned}$$

Also define a (0, 1)-*matrix* to be a matrix whose entries are all 0 or 1.

7.4.1 Proposition. *Let $\lambda \vdash n$, and let $\alpha = (\alpha_1, \alpha_2, \ldots)$ be a weak composition of n. Then the coefficient $M_{\lambda\alpha}$ of x^α in e_λ, i.e.,*

$$e_\lambda = \sum_{\mu\vdash n} M_{\lambda\mu} m_\mu, \tag{7.3}$$

is equal to the number of (0, 1)*-matrices $A = (a_{ij})_{i,j\geq 1}$ satisfying* $\text{row}(A) = \lambda$ *and* $\text{col}(A) = \alpha$.

Proof. Consider the matrix

$$X = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots \\ x_1 & x_2 & x_3 & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}.$$

To obtain a term of e_λ, choose λ_1 entries from the first row, λ_2 from the second row, etc. Let the product of the chosen entries be x^α. If we convert the chosen entries to 1's and the other entries to 0's, then we obtain a matrix A with $\text{row}(A) = \lambda$ and $\text{col}(A) = \alpha$. Conversely, any such matrix corresponds to a term of e_λ, and the proof follows. □

7.4.2 Corollary. *Let $M_{\lambda\mu}$ be given by (7.3). Then $M_{\lambda\mu} = M_{\mu\lambda}$, i.e., the transition matrix between the bases** $\{m_\lambda : \lambda \vdash n\}$ *and* $\{e_\lambda : \lambda \vdash n\}$ *is a symmetric matrix.*

Proof. The (0, 1)-matrix A satisfies $\text{row}(A) = \lambda$ and $\text{col}(A) = \mu$ if and only if the transpose A^t satisfies $\text{row}(A^t) = \mu$ and $\text{col}(A^t) = \lambda$. □

It is easy to see that the coefficient $M_{\lambda\mu}$ of (7.3) has the following alternative combinatorial interpretation. We have n balls in all, with λ_i balls labeled i. We also have boxes labeled $1, 2, \ldots$. Then $M_{\lambda\mu}$ is the number of ways of placing the balls into the boxes so that: (a) no box contains more than one ball with the same label, and (b) box i contains exactly μ_i balls. The elegant combinatorial interpretations we have given of $M_{\lambda\mu}$ are our first hints of the combinatorial efficacy of the theory of symmetric functions.

In general, let $\{u_\lambda\}$ be a basis for Λ and let $f \in \Lambda$. If the expansion $f = \sum_\lambda a_\lambda u_\lambda$ of f in terms of the basis $\{u_\lambda\}$ satisfies $a_\lambda \geq 0$ for all λ, then we say that f is *u-positive*. If f is u-positive, then the coefficients a_λ often have a simple combinatorial or algebraic interpretation. (An example of an algebraic interpretation would be the dimension of a vector space.) For instance, it is obvious from the relevant definitions that e_λ is m-positive, and Proposition 7.4.1 gives a stronger result (viz., a combinatorial interpretation of the coefficients). Similarly to the definition of u-positivity, we also say that f is *u-integral* if the coefficients a_λ above are integers.

Proposition 7.4.1 has an equivalent formulation in terms of generating functions. The type of generating function that we will be considering throughout this chapter has the form $z = \sum_\lambda c_\lambda u_\lambda$, where λ ranges over Par, $\{u_\lambda\}$ is a $\mathbb{Q}$-basis for Λ (and usually $u_\lambda \in \Lambda^n$, where $\lambda \vdash n$), and c_λ belongs to some coefficient ring R (which for us will always be a $\mathbb{Q}$-algebra). We may think of z as belonging to the
* ring $\hat{\Lambda}_R = \hat{\Lambda} \otimes R$, where $\hat{\Lambda}_R$ denotes the completion of Λ_R with respect to the ideal $\Lambda_R^1 \oplus \Lambda_R^2 \cdots$. Readers unfamiliar with completion need not be concerned; the generating functions $\sum c_\lambda u_\lambda$ will always behave in a reasonable, intuitive way.

A frequently occurring class of generating functions has $R = \Lambda(y)$, i.e., symmetric functions in a new set of variables $y = (y_1, y_2, \ldots)$. For instance, the generating function $z = \sum_\lambda m_\lambda(x) e_\lambda(y)$ of the next proposition is of the form $\sum_\lambda c_\lambda u_\lambda$ where $u_\lambda = m_\lambda(x) \in \Lambda(x)$ and $c_\lambda = e_\lambda(y) \in \Lambda(y) = R$. In general, if a function c_λ indexed by $\lambda \in$ Par arises in an enumeration problem, then it is

* It follows from Theorem 7.4.4 below that the set $\{e_\lambda : \lambda \vdash n\}$ is indeed a basis.

natural to consider a generating function of the form $\sum c_\lambda u_\lambda$ for a suitable basis u_λ of Λ. The difficulty, of course, is deciding what basis is "suitable." We will see throughout this chapter various choices of u_λ, the most common being the Schur functions s_λ discussed in Sections 7.10–7.19.

Let us now give the promised reformulation of Proposition 7.4.1. The generating function that we consider, as well as a closely related one to be discussed later (see Proposition 7.5.3), plays an important role in the theory of symmetric functions.

7.4.3 Proposition. *We have*

$$\prod_{i,j}(1+x_iy_j) = \sum_{\lambda,\mu} M_{\lambda\mu}m_\lambda(x)m_\mu(y) \tag{7.4}$$

$$= \sum_\lambda m_\lambda(x)e_\lambda(y). \tag{7.5}$$

Here λ and μ range over Par. *(It suffices to take $|\lambda| = |\mu|$, since otherwise $M_{\lambda\mu} = 0$.)*

Proof. A monomial $x_1^{\alpha_1}x_2^{\alpha_2}\cdots y_1^{\beta_1}y_2^{\beta_2}\cdots = x^\alpha y^\beta$ appearing in the expansion of $\prod(1 + x_iy_j)$ is obtained by choosing a (0, 1)-matrix $A = (a_{ij})$ with finitely many 1's, satisfying

$$\prod_{i,j}(x_iy_j)^{a_{ij}} = x^\alpha y^\beta.$$

But

$$\prod_{i,j}(x_iy_j)^{a_{ij}} = x^{\mathrm{row}(A)}y^{\mathrm{col}(A)},$$

so the coefficient of $x^\alpha y^\beta$ in the product $\prod(1 + x_iy_j)$ is the number of (0, 1)-matrices A satisfying $\mathrm{row}(A) = \alpha$ and $\mathrm{col}(A) = \beta$. Hence equation (7.4) follows. Equation (7.5) is then a consequence of (7.3). □

Note that Corollary 7.4.2 is immediate from (7.4), since the product $\prod(1+x_iy_j)$ is invariant under interchanging x_i and y_i for all i.

We now come to a basic result known as the "fundamental theorem of symmetric functions," though for us it will barely scratch the surface of this subject.

7.4.4 Theorem. *Let $\lambda, \mu \vdash n$. Then $M_{\lambda\mu} = 0$ unless $\mu \le \lambda'$, while $M_{\lambda\lambda'} = 1$. Hence the set $\{e_\lambda : \lambda \vdash n\}$ is a basis for Λ^n (so $\{e_\lambda : \lambda \in$ Par$\}$ is a basis for Λ). Equivalently, $e_1, e_2, \ldots$ are algebraically independent and generate Λ as a*

$\mathbb{Q}$-algebra, which we write as

$$\Lambda = \mathbb{Q}[e_1, e_2, \ldots].$$

Proof. Suppose that $M_{\lambda\mu} \neq 0$, so by Proposition 7.4.1 there is a (0, 1)-matrix A with $\mathrm{row}(A) = \lambda$ and $\mathrm{col}(A) = \mu$. Let A' be the matrix with $\mathrm{row}(A') = \lambda$ and with its 1's left-justified, i.e., $A'_{ij} = 1$ precisely for $1 \leq j \leq \lambda_i$. For any i the number of 1's in the first i columns of A' clearly is not less than the number of 1's in the first i columns of A, so by definition of dominance order we have $\mathrm{col}(A') \geq \mathrm{col}(A) = \mu$. But $\mathrm{col}(A') = \lambda'$, so $\lambda' \geq \mu$ as desired. Moreover, it is easy to see that A' is the *only* (0, 1)-matrix with $\mathrm{row}(A') = \lambda$ and $\mathrm{col}(A') = \lambda'$, so $M_{\lambda\lambda'} = 1$.

The previous argument shows the following: let $\lambda^1, \lambda^2, \ldots, \lambda^{p(n)}$ be an ordering of $\mathrm{Par}(n)$ that is compatible with dominance order, and such that the "reverse conjugate" order $(\lambda^{p(n)})', \ldots, (\lambda^2)', (\lambda^1)'$ is also compatible with dominance order. (It is easily seen that such orders exist.) Then the matrix $(M_{\lambda\mu})$, with the row order $\lambda^1, \lambda^2, \ldots$ and column order $(\lambda^1)', (\lambda^2)', \ldots$, is upper triangular with 1's on the main diagonal. Hence it is invertible, so $\{e_\lambda : \lambda \vdash n\}$ is a basis for Λ^n. (In fact, it is a basis for $\Lambda^n_{\mathbb{Z}}$ since the diagonal entries are actually 1's, and not merely nonzero.)

The set $\{e_\lambda : \lambda \in \mathrm{Par}\}$ consists of all monomials $e_1^{a_1} e_2^{a_2} \cdots$ (where $a_i \in \mathbb{N}$, $\sum a_i < \infty$). Hence the linear independence of $\{e_\lambda : \lambda \in \mathrm{Par}\}$ is equivalent to the algebraic independence of $e_1, e_2, \ldots$, as desired. $\square$

Figure 7-2 gives a short table of the coefficients $M_{\lambda\mu}$.

$$\begin{array}{lcl}
e_1 & = & m_1 \\
\\
e_{11} & = & m_2 + 2m_{11} \\
e_2 & = & m_{11} \\
\\
e_{111} & = & m_3 + 3m_{21} + 6m_{111} \\
e_{21} & = & m_{21} + 3m_{111} \\
e_3 & = & m_{111} \\
\\
e_{1111} & = & m_4 + 4m_{31} + 6m_{22} + 12m_{211} + 24m_{1111} \\
e_{211} & = & m_{31} + 2m_{22} + 5m_{211} + 12m_{1111} \\
e_{22} & = & m_{22} + 2m_{211} + 6m_{1111} \\
e_{31} & = & m_{211} + 4m_{1111} \\
e_4 & = & m_{1111} \\
\\
e_{11111} & = & m_5 + 5m_{41} + 10m_{32} + 20m_{311} + 30m_{221} + 60m_{2111} + 120m_{11111} \\
e_{2111} & = & m_{41} + 3m_{32} + 7m_{311} + 12m_{221} + 27m_{2111} + 60m_{11111} \\
e_{221} & = & m_{32} + 2m_{311} + 5m_{221} + 12m_{2111} + 30m_{11111} \\
e_{311} & = & m_{311} + 2m_{221} + 7m_{2111} + 20m_{11111} \\
e_{32} & = & m_{221} + 3m_{2111} + 10m_{11111} \\
e_{41} & = & m_{2111} + 5m_{11111} \\
e_5 & = & m_{11111}
\end{array}$$

Figure 7-2. The coefficients $M_{\lambda\mu}$.

7.5 Complete Homogeneous Symmetric Functions

Define the *complete homogeneous symmetric functions* (or just *complete symmetric functions*) h_λ for $\lambda \in$ Par by the formulas

$$\begin{aligned} h_n &= \sum_{\lambda \vdash n} m_\lambda = \sum_{i_1 \le \cdots \le i_n} x_{i_1} \cdots x_{i_n}, \quad n \ge 1 \qquad (\text{with } h_0 = m_\emptyset = 1) \\ h_\lambda &= h_{\lambda_1} h_{\lambda_2} \cdots \qquad \text{if } \lambda = (\lambda_1, \lambda_2, \ldots). \end{aligned} \tag{7.6}$$

Thus h_n is the sum of all monomials of degree n.

The symmetric functions h_λ are in many ways "dual" to the elementary symmetric functions e_μ. The underlying reason for this duality will be brought out by Theorem 7.6.1 and various subsequent developments. For now let us consider the "complete analogue" of Proposition 7.4.1.

7.5.1 Proposition. *Let $\lambda \vdash n$, and let $\alpha = (\alpha_1, \alpha_2, \ldots)$ be a weak composition of n. Then the coefficient $N_{\lambda\alpha}$ of x^α in h_λ, i.e.,*

$$h_\lambda = \sum_{\mu \vdash n} N_{\lambda\mu} m_\mu, \tag{7.7}$$

is equal to the number of $\mathbb{N}$-matrices $A = (a_{ij})_{i,j\ge1}$ satisfying $\mathrm{row}(A) = \lambda$ *and* $\mathrm{col}(A) = \alpha$.

Proof. Analogous to the proof of Proposition 7.4.1. A term x^α from $h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots$ is obtained by choosing a term $x_1^{a_{i1}} x_2^{a_{i2}} \cdots$ from each h_{λ_i} such that

$$\prod_i \left(x_1^{a_{i1}} x_2^{a_{i2}} \cdots \right) = x^\alpha.$$

But this is just the same as choosing (a_{ij}) to be an $\mathbb{N}$-matrix A with $\mathrm{row}(A) = \lambda$ and $\mathrm{col}(A) = \alpha$, and the proof follows. □

7.5.2 Corollary. *Let $N_{\lambda\mu}$ be given by (7.7). Then $N_{\lambda\mu} = N_{\mu\lambda}$, i.e., the transition matrix between the bases* $\{m_\lambda : \lambda \vdash n\}$ and $\{h_\lambda : \lambda \vdash n\}$ is a symmetric matrix.*

Proof. Exactly analogous to the proof of Corollary 7.4.2. □

The coefficient $N_{\lambda\mu}$ of (7.7) has an alternative combinatorial interpretation in terms of balls into boxes, similar to the interpretation of $M_{\lambda\mu}$ in this way. We have

* It follows from Corollary 7.6.2 that the set $\{h_\lambda : \lambda \vdash n\}$ is indeed a basis.

n balls in all, with λ_i balls labeled i. We also have boxes labeled $1, 2, \ldots$. Then $N_{\lambda\mu}$ is the number of ways of placing the balls in the boxes so that box i contains exactly μ_i balls.

On the combinatorial level, the duality between e_λ and h_μ is manifested by $(0, 1)$-matrices *vs.* $\mathbb{N}$-matrices, or equivalently, balls into boxes (subject to certain conditions) not allowing repetitions or allowing repetitions. This situation is reminiscent of the reciprocity between $\binom{n}{k}$ and $\left(\!\binom{n}{k}\!\right) = (-1)^k\binom{-n}{k}$, or of the more general reciprocity theorems of Sections 4.5 and 4.6. Indeed, given a symmetric function $f(x)$ and $n \in \mathbb{N}$, let us write

$$f(1^n) = f(x_1 = x_2 = \cdots = x_n = 1,\ x_{n+1} = x_{n+2} = \cdots = 0). \tag{7.8}$$

Then

$$\begin{aligned} e_k(1^n) &= \sum_{1 \le i_1 < \cdots < i_k \le n} 1 = \binom{n}{k} \\ h_k(1^n) &= \sum_{1 \le i_1 \le \cdots \le i_k \le n} 1 = \left(\!\binom{n}{k}\!\right). \end{aligned}$$

We next give the generating function interpretation of Proposition 7.5.1. The proof is analogous to that of Proposition 7.4.3 and is omitted.

$$\begin{aligned}
h_1 &= m_1 \\[1ex]
h_{11} &= 2m_{11} + m_2 \\
h_2 &= m_{11} + m_2 \\[1ex]
h_{111} &= 6m_{111} + 3m_{21} + m_3 \\
h_{21} &= 3m_{111} + 2m_{21} + m_3 \\
h_3 &= m_{111} + m_{21} + m_3 \\[1ex]
h_{1111} &= 24m_{1111} + 12m_{211} + 6m_{22} + 4m_{31} + m_4 \\
h_{211} &= 12m_{1111} + 7m_{211} + 4m_{22} + 3m_{31} + m_4 \\
h_{22} &= 6m_{1111} + 4m_{211} + 3m_{22} + 2m_{31} + m_4 \\
h_{31} &= 4m_{1111} + 3m_{211} + 2m_{22} + 2m_{31} + m_4 \\
h_4 &= m_{1111} + m_{211} + m_{22} + m_{31} + m_4 \\[1ex]
h_{11111} &= 120m_{11111} + 60m_{2111} + 30m_{221} + 20m_{311} + 10m_{32} + 5m_{41} + m_5 \\
h_{2111} &= 60m_{11111} + 33m_{2111} + 18m_{221} + 13m_{311} + 7m_{32} + 4m_{41} + m_5 \\
h_{221} &= 30m_{11111} + 18m_{2111} + 11m_{221} + 8m_{311} + 5m_{32} + 3m_{41} + m_5 \\
h_{311} &= 20m_{11111} + 13m_{2111} + 8m_{221} + 7m_{311} + 4m_{32} + 3m_{41} + m_5 \\
h_{32} &= 10m_{11111} + 7m_{2111} + 5m_{221} + 4m_{311} + 3m_{32} + 2m_{41} + m_5 \\
h_{41} &= 5m_{11111} + 4m_{2111} + 3m_{221} + 3m_{311} + 2m_{32} + m_{41} + m_5 \\
h_5 &= m_{11111} + m_{2111} + m_{221} + m_{311} + m_{32} + m_{41} + m_5
\end{aligned}$$

* **Figure 7-3.** The coefficients $N_{\lambda\mu}$.

7.5.3 Proposition. *We have*

$$\prod_{i,j}(1 - x_i y_j)^{-1} = \sum_{\lambda,\mu} N_{\lambda\mu} m_\lambda(x) m_\mu(y) \tag{7.9}$$

$$= \sum_{\lambda} m_\lambda(x) h_\lambda(y), \tag{7.10}$$

where λ and μ range over Par *(and where it suffices to take $|\lambda| = |\mu|$).*

It is not as easy to argue as in the proof of Theorem 7.4.4 that the set $\{h_\lambda : \lambda \vdash n\}$ is a basis for Λ^n (so $h_1, h_2, \ldots$ are algebraically independent), since the matrix $(N_{\lambda\mu})$ has no nice triangularity properties. But the linear independence of the h_λ's will be a trivial consequence of Theorem 7.6.1, so we save its "official" statement until then.

Figure 7-3 gives a short table of the coefficients $N_{\lambda\mu}$.

7.6 An Involution

Since $\Lambda = \mathbb{Q}[e_1, e_2, \ldots]$, an algebra endomorphism $f : \Lambda \to \Lambda$ is determined uniquely by its values $f(e_n)$, $n \geq 1$; and conversely any choice of $f(e_n) \in \Lambda$ determines an endomorphism f. Define an endomorphism $\omega : \Lambda \to \Lambda$ by $\omega(e_n) = h_n$, $n \geq 1$. Thus (since ω preserves multiplication) $\omega(e_\lambda) = h_\lambda$ for all partitions λ.

7.6.1 Theorem. *The endomorphism ω is an involution, i.e., $\omega^2 = 1$ (the identity automorphism), or equivalently $\omega(h_n) = e_n$. (Thus $\omega(h_\lambda) = e_\lambda$ for all partitions λ.)*

Proof. Consider the formal power series

$$H(t) := \sum_{n \geq 0} h_n t^n \in \Lambda[[t]]$$

$$E(t) := \sum_{n \geq 0} e_n t^n \in \Lambda[[t]].$$

We leave to the reader the easy verification of the identities

$$H(t) = \prod_n (1 - x_n t)^{-1} \tag{7.11}$$

$$E(t) = \prod_n (1 + x_n t). \tag{7.12}$$

Hence $H(t)E(-t) = 1$. Equating coefficients of t^n on both sides yields

$$0 = \sum_{i=0}^{n} (-1)^i e_i h_{n-i}, \qquad n \geq 1. \tag{7.13}$$

Conversely, if $\sum_{i=0}^{n}(-1)^i u_i h_{n-i} = 0$ for all $n \geq 1$, for certain $u_i \in \Lambda$ with $u_0 = 1$,

then $u_i = e_i$. Now apply ω to (7.13) to obtain

$$0 = \sum_{i=0}^{n}(-1)^i h_i \omega(h_{n-i}) = (-1)^n \sum_{i=0}^{n}(-1)^i \omega(h_i) h_{n-i} = 0,$$

whence $\omega(h_i) = e_i$ as desired. □

The involution ω may be regarded as an algebraic elaboration of the reciprocity between sets and multisets expressed by the identity $\left(\!\binom{n}{k}\!\right) = (-1)^k \binom{-n}{k}$, as suggested in Section 7.5.

7.6.2 Corollary. *The set $\{h_\lambda : \lambda \vdash n\}$ is a basis for Λ^n (so $\{h_\lambda : \lambda \in \mathrm{Par}\}$ is a basis for Λ). Equivalently, $h_1, h_2, \ldots$ are algebraically independent and generate Λ as a $\mathbb{Q}$-algebra, which we write as*

$$\Lambda = \mathbb{Q}[h_1, h_2, \ldots].$$

Proof. Theorem 7.6.1 shows that the endomorphism $\omega : \Lambda \to \Lambda$ defined by $\omega(e_n) = h_n$ is invertible (since $\omega = \omega^{-1}$), and hence is an automorphism of Λ. The proof now follows from Theorem 7.4.4. □

7.7 Power Sum Symmetric Functions

We define a fourth set p_λ of symmetric functions indexed by $\lambda \in \mathrm{Par}$ and called *power sum symmetric functions*, as follows:

$$p_n = m_n = \sum_i x_i^n, \quad n \geq 1 \qquad (\text{with } p_0 = m_\emptyset = 1)$$

$$p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots \qquad \text{if } \lambda = (\lambda_1, \lambda_2, \ldots).$$

7.7.1 Proposition. *Let $\lambda = (\lambda_1, \ldots, \lambda_\ell) \vdash n$, where $\ell = \ell(\lambda)$, and set*

$$p_\lambda = \sum_{\mu \vdash n} R_{\lambda\mu} m_\mu. \tag{7.14}$$

Let $k = \ell(\mu)$. Then $R_{\lambda\mu}$ is equal to the number of ordered partitions $\pi = (B_1, \ldots, B_k)$ of the set $[\ell]$ such that

$$\mu_j = \sum_{i \in B_j} \lambda_i, \qquad 1 \leq j \leq k. \tag{7.15}$$

Proof. $R_{\lambda\mu}$ is the coefficient of $x^\mu = x_1^{\mu_1} x_2^{\mu_2} \cdots$ in $p_\lambda = (\sum x_i^{\lambda_1})(\sum x_i^{\lambda_2}) \cdots$. To obtain the monomial x^μ in the expansion of this product, we choose a term $x_{i_j}^{\lambda_j}$ from each factor $\sum x_i^{\lambda_j}$ so that $\prod_j x_{i_j}^{\lambda_j} = x^\mu$. Define $B_r = \{j : i_j = r\}$. Then $(B_1, \ldots, B_k)$ will be an ordered partition of $[\ell]$ satisfying (7.15), and conversely every such ordered partition gives rise to a term x^μ. □

7.7.2 Corollary. *Let $R_{\lambda\mu}$ be as in (7.14). Then $R_{\lambda\mu} = 0$ unless $\lambda \leq \mu$, while*

$$R_{\lambda\lambda} = \prod_i m_i(\lambda)!, \tag{7.16}$$

where $\lambda = \langle 1^{m_1(\lambda)} 2^{m_2(\lambda)} \cdots \rangle$. Hence $\{p_\lambda : \lambda \vdash n\}$ is a basis for Λ^n (so $\{p_\lambda : \lambda \in$ Par$\}$ is a basis for Λ). Equivalently, $p_1, p_2, \ldots$ are algebraically independent and generate Λ as a $\mathbb{Q}$-algebra, i.e.,

$$\Lambda = \mathbb{Q}[p_1, p_2, \ldots].$$

Proof. If $R_{\lambda\mu} \neq 0$, then by Proposition 7.7.1 there is an ordered partition $\pi = (B_1, \ldots, B_k)$ of the set $[\ell] = [\ell(\lambda)]$ satisfying (7.15). Given $1 \leq r \leq \ell$, let $B_{i_1}, \ldots, B_{i_s}$ be the *distinct* blocks of π containing at least one of $1, 2, \ldots, r$. From (7.15) we have $\mu_{i_1} + \cdots + \mu_{i_s} \geq \lambda_1 + \cdots + \lambda_r$. But $\mu_1 + \cdots + \mu_r \geq \mu_{i_1} + \cdots + \mu_{i_s}$ since $r \geq s$ and $\mu_1 \geq \mu_2 \geq \cdots$. Hence $\mu \geq \lambda$, as desired.

If $\mu = \lambda$, then each block B_i is a singleton $\{j\}$, which we denote simply as j. $B_1, \ldots, B_{m_1}$ can be any ordering of $1, \ldots, m_1$. Then $B_{m_1+1}, \ldots, B_{m_1+m_2}$ can be any ordering of $m_1 + 1, \ldots, m_1 + m_2$, etc., giving a total of $R_{\lambda\lambda} = m_1! m_2! \cdots$ possibilities for π.

The fact that $\{p_\lambda : \lambda \vdash n\}$ is a basis for Λ (so $\Lambda = \mathbb{Q}[p_1, p_2, \ldots]$) follows by reasoning as in the proof of Theorem 7.4.4. □

NOTE. Because the diagonal elements $R_{\lambda\lambda}$ are not all ± 1, it follows that $\{p_\lambda : \lambda \vdash n\}$ is not a $\mathbb{Z}$-basis for $\Lambda^n_{\mathbb{Z}}$. Rather, the (additive) abelian subgroup P_n of $\Lambda^n_{\mathbb{Z}}$ generated by the p_λ's has index

$$\begin{aligned}\left[\Lambda^n_{\mathbb{Z}} : P_n\right] &= \det(R_{\lambda\mu})_{\lambda,\mu\vdash n} \\ &= \prod_{\lambda\vdash n} \prod_i m_i(\lambda)!.\end{aligned}$$

By Exercise 1.26, if follows that this index is also given by

$$\left[\Lambda^n_{\mathbb{Z}} : P_n\right] = \prod_{\lambda\vdash n} 1^{m_1(\lambda)} 2^{m_2(\lambda)} \cdots,$$

i.e., the product of all parts of all partitions of n.

We now consider the effect of the involution ω on p_λ. A generating function approach is most efficacious. For any partition $\lambda = \langle 1^{m_1} 2^{m_2} \cdots \rangle$, define

$$z_\lambda = 1^{m_1} m_1! \, 2^{m_2} m_2! \cdots. \tag{7.17}$$

For instance, $z_{442111} = 1^3 3! \, 2^1 1! \, 4^2 2! = 384$. If $w \in \mathfrak{S}_n$, then the *cycle type* $\rho(w)$ of w is the partition $\rho(w) = (\rho_1, \rho_2, \ldots) \vdash n$ such that the cycle lengths of w (in its factorization into disjoint cycles) are $\rho_1, \rho_2, \ldots$. Recall from Proposition 1.3.2 that the number of permutations $w \in \mathfrak{S}_n$ of a fixed cycle type $\rho = \langle 1^{m_1} 2^{m_2} \cdots \rangle$

is given by

$$\begin{aligned}\#\{w \in \mathfrak{S}_n : \rho(w) = \rho\} &= \frac{n!}{1^{m_1} m_1!\, 2^{m_2} m_2! \cdots} \\ &= n!\, z_\rho^{-1}. \end{aligned} \tag{7.18}$$

The set $\{v \in \mathfrak{S}_n : \rho(v) = \rho\}$ is just the conjugacy class in $\mathfrak{S}_n$ containing w. For any finite group G, the order $\#K_w$ of the conjugacy class K_w containing w is equal to the index $[G : C(w)]$ of the centralizer of w. Hence:

7.7.3 Proposition. *Let $\lambda \vdash n$. Then z_λ is equal to the number of permutations $v \in \mathfrak{S}_n$ that commute with a fixed w_λ of cycle type λ.*

For a bijective proof of Proposition 7.7.3, see Exercise 7.6.
For a partition $\lambda = \langle 1^{m_1} 2^{m_2} \cdots \rangle$ of n, define

$$\varepsilon_\lambda = (-1)^{m_2+m_4+\cdots} = (-1)^{n-\ell(\lambda)}. \tag{7.19}$$

Thus for $w \in \mathfrak{S}_n$, $\varepsilon_{\rho(w)}$ is $+1$ if w is an even permutation and -1 otherwise, so the map $\mathfrak{S}_n \to \{\pm 1\}$ defined by $w \mapsto \varepsilon_{\rho(w)}$ is the usual "sign homomorphism."

7.7.4 Proposition. *We have*

$$\begin{aligned}\prod_{i,j}(1 - x_i y_j)^{-1} &= \exp \sum_{n\geq 1} \frac{1}{n} p_n(x) p_n(y) \\ &= \sum_\lambda z_\lambda^{-1} p_\lambda(x) p_\lambda(y) \end{aligned} \tag{7.20}$$

$$\begin{aligned}\prod_{i,j}(1 + x_i y_j) &= \exp \sum_{n\geq 1} \frac{1}{n} (-1)^{n-1} p_n(x) p_n(y) \\ &= \sum_\lambda z_\lambda^{-1} \varepsilon_\lambda\, p_\lambda(x) p_\lambda(y). \end{aligned} \tag{7.21}$$

Proof. We have

$$\begin{aligned}\log \prod_{i,j}(1 - x_i y_j)^{-1} &= \sum_{i,j} \log(1 - x_i y_j)^{-1} \\ &= \sum_{i,j} \sum_{n\geq 1} \frac{1}{n} x_i^n y_j^n \\ &= \sum_{n\geq 1} \frac{1}{n} \left(\sum_i x_i^n\right) \left(\sum_j y_j^n\right) \\ &= \sum_{n\geq 1} \frac{1}{n} p_n(x) p_n(y). \end{aligned}$$

Thus the first equality of (7.20) is established. The second equality of (7.20) is a consequence of (and in fact is equivalent to) the permutation version of the exponential formula (Corollary 5.1.9). More specifically, in Corollary 5.1.9 set $f(n) \doteq p_n(x)p_n(y)$ and $x = 1$. Then

$$\begin{aligned} h(n) &= \sum_{w \in \mathfrak{S}_n} p_{\rho(w)}(x)p_{\rho(w)}(y) \\ &= \sum_{\lambda \vdash n} c_\lambda p_\lambda(x)p_\lambda(y), \end{aligned}$$

where $c_\lambda = \#\{w \in \mathfrak{S}_n : \rho(w) = \lambda\}$. By equation (7.18) we have $c_\lambda = n!z_\lambda^{-1}$, and (7.20) follows from Corollary 5.1.9. The proof of (7.21) is entirely analogous. □

From the previous proposition it is easy to deduce the effect of ω on p_λ.

7.7.5 Proposition. *Let $\lambda \vdash n$. Then*

$$\omega p_\lambda = \varepsilon_\lambda p_\lambda.$$

In other words, p_λ is an eigenvector for ω corresponding to the eigenvalue ε_λ.

Proof. Regard ω as acting on symmetric functions in the variables $y = (y_1, y_2, \ldots)$; those in the variables x are regarded as scalars. Apply ω to (7.20). We obtain

$$\begin{aligned} \omega \sum_\lambda z_\lambda^{-1} p_\lambda(x)p_\lambda(y) &= \omega \prod_{i,j} (1 - x_i y_j)^{-1} \\ &= \sum_\nu m_\nu(x)\omega h_\nu(y) \qquad \text{(by (7.10))} \\ &= \sum_\nu m_\nu(x)e_\nu(y) \qquad \text{(by Theorem 7.6.1)} \\ &= \prod_{i,j}(1 + x_i y_j) \qquad \text{(by (7.5))} \\ &= \sum_\lambda z_\lambda^{-1}\varepsilon_\lambda\, p_\lambda(x)p_\lambda(y) \qquad \text{(by (7.21)).} \end{aligned}$$

Since the $p_\lambda(x)$'s are linearly independent, their coefficients in the first and last sums of the above chain of equalities must be the same. In other words, $\omega p_\lambda(y) = \varepsilon_\lambda p_\lambda(y)$, as desired. □

Note in particular that $\omega p_n = (-1)^{n-1} p_n$, or $\omega p_n(x) = -p_n(-x)$. Since ω is an automorphism, just the values of ωp_n in fact suffice to determine ωp_λ for any partition λ.

Consider ω restricted to the vector space (of dimension $p(n)$) Λ^n. Since the p_λ's for $\lambda \vdash n$ are linearly independent, it follows from Proposition 7.7.5 that the characteristic polynomial (normalized to be monic) of the linear transformation $\omega : \Lambda^n \to \Lambda^n$ is given by $(x-1)^{e(n)}(x+1)^{o(n)}$, where $e(n)$ (respectively, $o(n)$) is the number of partitions of n with an even number (respectively, odd number) of even parts. In other words, $e(n)$ is the number of *even conjugacy classes* (i.e., conjugacy classes contained in the alternating group) of $\mathfrak{S}_n$. By Exercise 1.1.9(b) we have

$$(x-1)^{e(n)}(x+1)^{o(n)} = (x^2-1)^{o(n)}(x-1)^{k(n)},$$

where $k(n)$ is the number of self-conjugate partitions of n. At the end of Section 7.14 we will see a simple reason for the factor $(x-1)^{k(n)}$ in terms of symmetric functions.

We can now ask how to express the symmetric functions m_λ, h_λ, and e_λ in terms of the p_μ's. Although combinatorial interpretations can be given to the coefficients in these expansions, they tend to be messy and not very useful. One special case, however, is of considerable importance.

7.7.6 Proposition. *We have*

$$h_n = \sum_{\lambda \vdash n} z_\lambda^{-1} p_\lambda \tag{7.22}$$

$$e_n = \sum_{\lambda \vdash n} \varepsilon_\lambda z_\lambda^{-1} p_\lambda. \tag{7.23}$$

Proof. Substituting $y = (t, 0, 0, \ldots)$ in (7.20) immediately yields (7.22). Equation (7.23) is similarly obtained from (7.21), or by applying ω to (7.22). $\square$

See Example 5.2.11 for a combinatorial proof of equation (7.22). Equation (7.23) can be given a similar proof.

7.8 Specializations

In many combinatorial problems involving symmetric functions f we only need partial information about f, such as a particular coefficient or value. In this section we give a brief overview of the most common specializations that arise in practice. (See Exercises 7.43 and 7.44 for two others.) Proofs for the most part are straightforward and will be omitted. First let us give a formal definition of the concept of specialization.

7.8.1 Definition. Let R be a commutative $\mathbb{Q}$-algebra with identity. A *specialization* of the ring Λ is a homomorphism $\varphi : \Lambda \to R$. (We always assume homomorphisms are unital, i.e., $\varphi(1) = 1$.)

The most obvious examples of specializations arise from substituting elements a_i of R for the variables x_i (provided of course this substitution is well-defined formally; it would make no sense, for instance, to set each $x_i = 1$ in $h_1(x) = x_1 + x_2 + \cdots$). We may then write

$$\varphi(f) = f(a_1, a_2, \ldots),$$

and we call φ the *substitution* of a_i for x_i.

Our first example may be called "reducing the number of variables" and is very common. Let Λ_n denote the set of all polynomials $f \in \mathbb{Q}[x_1, \ldots, x_n]$ in the variables $x_1, \ldots, x_n$ with rational coefficients which are invariant under any permutation of the variables. Thus f is just a symmetric function in the variables $x_1, \ldots, x_n$. Define $r_n : \Lambda \to \Lambda_n$ by $r_n(f) = f(x_1, \ldots, x_n, 0, 0, \ldots)$ (written $f(x_1, \ldots, x_n)$). The next proposition examines the behavior of the four bases m_λ, p_λ, h_λ, e_λ, as well as the involution ω, under r_n.

7.8.2 Proposition. *Let* Par_n *denote the set of all partitions* $\lambda \in \mathrm{Par}$ *of length at most* n*, i.e.,*

$$\mathrm{Par}_n = \{\lambda \in \mathrm{Par} : \ell(\lambda) \le n\}.$$

(a) *The sets* $\{r_n(m_\lambda) : \lambda \in \mathrm{Par}_n\}$, $\{r_n(p_\lambda) : \lambda' \in \mathrm{Par}_n\}$, $\{r_n(h_\lambda) : \lambda' \in \mathrm{Par}_n\}$, $\{r_n(e_\lambda) : \lambda' \in \mathrm{Par}_n\}$ *are all* $\mathbb{Q}$*-bases for* Λ_n*. Moreover, if* $\lambda \notin \mathrm{Par}_n$*, then* $r_n(m_\lambda) = r_n(e_{\lambda'}) = 0$.

(b) *For convenience identify an element* $f \in \Lambda$ *with its image* $r_n(f)$ *in* Λ_n*. Define a linear transformation* $\omega_n : \Lambda_n \to \Lambda_n$ *by* $\omega_n(e_\lambda) = h_\lambda$ *for* $\lambda' \in \mathrm{Par}_n$*. Then* ω_n *is an algebra automorphism and an involution, and* $\omega_n(p_\lambda) = \varepsilon_\lambda p_\lambda$ *for* $\lambda' \in \mathrm{Par}_n$.

Proof. Straightforward consequence of analogous properties for Λ and triangularity properties of the bases m_λ, p_λ, e_λ discussed previously. □

A little caution is needed when dealing with p_λ or h_λ in Λ_n when $\lambda' \notin \mathrm{Par}_n$. For instance, p_λ need not be zero nor an eigenvector of ω_n. For instance, when $n = 2$ we have $p_3 = \frac{1}{2}(3p_{21} - p_{111})$ and $\omega_2(p_3) = \frac{1}{2}(-3p_{21} - p_{111})$.

An important substitution $\mathrm{ps}_n : \Lambda \to \mathbb{Q}[q]$ is defined by

$$\mathrm{ps}_n(f) = f(1, q, q^2, \ldots, q^{n-1}),$$

and is called the *principal specialization* (of order n) of f. If we let $n \to \infty$ we obtain the limiting value

$$\mathrm{ps}(f) = f(1, q, q^2, \ldots) \in \mathbb{Q}[[q]], \tag{7.24}$$

called the *stable principal specialization* of f. (It is easily seen that $\lim_{n\to\infty} \mathrm{ps}_n(f)$ exists in the sense of Section 1.1.) A specialization $\mathrm{ps}_n^1 : \Lambda \to \mathbb{Q}$ of the principal

specialization is obtained by letting $q = 1$, i.e.,

$$\mathrm{ps}_n^1(f) = f(\underbrace{1, 1, \ldots, 1}_{n \text{ 1's}}) = f(1^n),$$

in the notation of (7.8).

7.8.3 Proposition. *The following table summarizes the behavior of the bases* $m_\lambda, p_\lambda, h_\lambda, e_\lambda$ *under* ps_n, ps, *and* ps_n^1:

basis b_λ	$\mathrm{ps}_n(b_\lambda)$	$\mathrm{ps}(b_\lambda)$	$\mathrm{ps}_n^1(b_\lambda)$
m_λ	messy	messy	$\binom{n}{\ell(\lambda)}\binom{\ell(\lambda)}{m_1(\lambda), m_2(\lambda), \ldots}$
p_λ	$\prod_{i=1}^{\ell} \frac{1-q^{n\lambda_i}}{1-q^{\lambda_i}}$	$\prod_{i=1}^{\ell} \frac{1}{1-q^{\lambda_i}}$	$n^{\ell(\lambda)}$
e_λ	$\prod_i q^{\binom{\lambda_i}{2}}\binom{n}{\lambda_i}$	$\prod_i \frac{q^{\binom{\lambda_i}{2}}}{(1-q)(1-q^2)\cdots(1-q^{\lambda_i})}$	$\prod_i \binom{n}{\lambda_i}$
h_λ	$\prod_i \binom{n+\lambda_i-1}{\lambda_i}$	$\prod_i \frac{1}{(1-q)(1-q^2)\cdots(1-q^{\lambda_i})}$	$\prod_i \left(\!\binom{n}{\lambda_i}\!\right)$

Proof. All can be done by straightforward combinatorial or algebraic reasoning. As an example, we show how to obtain $\mathrm{ps}_n(h_\lambda)$.

Since $h_\lambda = h_{\lambda_1} h_{\lambda_2} \cdots$ and ps_n is an algebra homomorphism, it suffices to compute $\mathrm{ps}_n(h_k)$. Since

$$h_k = \sum_{a_1+a_2+\cdots=k} x_1^{a_1} x_2^{a_2} \cdots,$$

we have

$$\mathrm{ps}_n(h_k) = \sum_{a_1+\cdots+a_n=k} q^{a_2+2a_3+\cdots+(n-1)a_n},$$

summed over all weak compositions of k into n parts. If we identify the sequence $(a_1, a_2, \ldots, a_n)$ with the partition $\lambda = \langle 1^{a_2}, 2^{a_3}, \ldots, (n-1)^{a_n}\rangle$, then we see that

$$\mathrm{ps}_n(h_k) = \sum_{\substack{\ell(\lambda)\le k \\ \ell(\lambda')\le n-1}} q^{|\lambda|} = \binom{n+k-1}{k},$$

by Proposition 1.3.19, and the proof follows. $\square$

We now consider an important specialization that is not obtained simply by substituting for the variables x_i. We call this specialization the *exponential specialization* $\mathrm{ex} : \Lambda \to \mathbb{Q}[t]$ or $\mathrm{ex} : \hat{\Lambda} \to \mathbb{Q}[[t]]$, defined by

$$\mathrm{ex}(p_n) = t\delta_{1n}.$$

Note that since the p_n's are algebraically independent and generate Λ as a $\mathbb{Q}$-algebra, any homomorphism $\varphi : \Lambda \to R$ is determined by its values $\varphi(p_n)$. Here we are setting $\varphi(p_1) = t$ and $\varphi(p_n) = 0$ if $n > 1$. (If the domain of ex is taken to be $\hat{\Lambda}$, then we define ex to preserve infinite linear combinations, or equivalently, to be continuous in a suitable topology.)

7.8.4 Proposition. (a) *We have*

$$\mathrm{ex}(f) = \sum_{n\geq 0}[x_1x_2\cdots x_n]f\frac{t^n}{n!}, \tag{7.25}$$

for any $f \in \hat{\Lambda}$, where $[x_1x_2\cdots x_n]f$ denotes the coefficient of $x_1x_2\cdots x_n$ in f. Equivalently, if $f = \sum_\lambda c_\lambda m_\lambda$, then

$$\mathrm{ex}(f) = \sum_{n\geq 0} c_{1^n}\frac{t^n}{n!}.$$

(b) *We have*

$$\mathrm{ex}(m_\lambda) = \begin{cases} \dfrac{t^n}{n!} & \text{if } \lambda = \langle 1^n \rangle \\ 0, & \text{otherwise} \end{cases}$$

$$\mathrm{ex}(p_\lambda) = \begin{cases} t^n & \text{if } \lambda = \langle 1^n \rangle \\ 0, & \text{otherwise} \end{cases} \tag{7.26}$$

$$\mathrm{ex}(h_\lambda) = \mathrm{ex}(e_\lambda) = \frac{t^{|\lambda|}}{\lambda_1!\lambda_2!\cdots}. \tag{7.27}$$

Proof. (a) Since the right-hand side of (7.25) is linear in f, we need only verify (7.25) for $f = p_\lambda$. This is a routine computation.
(b) Easy consequence of (a). □

7.8.5 Example. Let

$$F(x) = \prod_i (1 - x_i)^{-1} \cdot \prod_{i<j}(1 - x_ix_j)^{-1}.$$

(For the significance of this product, see Corollary 7.13.8.) In this example we will

evaluate $\mathrm{ex}(F(x))$. We can save a little work by observing that

$$\mathrm{ex}\prod_i (1-x_i)^{-1} = \mathrm{ex}\sum_{n\geq 0} h_n$$
$$= \sum_{n\geq 0}\frac{t^n}{n!} = e^t,$$

so (since ex is a homomorphism),

$$\mathrm{ex}(F(x)) = e^t \cdot \mathrm{ex}\prod_{i<j}(1-x_ix_j)^{-1}.$$

Now

$$\prod_{i<j}(1-x_ix_j)^{-1} = \exp\sum_{i<j}\log(1-x_ix_j)^{-1}$$
$$= \exp\sum_{i<j}\sum_{n\geq 1}\frac{(x_ix_j)^n}{n}$$
$$= \exp\sum_{n\geq 1}\frac{1}{2n}\left(p_n^2 - p_{2n}\right).$$

(Do not confuse ex with the ordinary exponential function exp.) Hence

$$\mathrm{ex}\prod_{i<j}(1-x_ix_j)^{-1} = \exp\sum_{n\geq 1}\frac{1}{2n}\mathrm{ex}\left(p_n^2 - p_{2n}\right)$$
$$= e^{t^2/2},$$

by (7.26), so

$$\mathrm{ex}(F(x)) = e^{t+\frac{1}{2}t^2}.$$

We recognize from equation (5.32) that $e^{t+\frac{1}{2}t^2}$ is the exponential generating function for the number $e_2(n)$ of involutions in $\mathfrak{S}_n$. Indeed, it is easy to see directly from the definition of $F(x)$ that

$$[x_1\cdots x_n]F(x) = e_2(n).$$

In view of Proposition 7.8.4, it is natural to ask whether the specialization ex has a "natural" q-analogue ex_q. The definition that works best is given by

$$\mathrm{ex}_q(h_n) = \frac{t^n}{(\boldsymbol{n})!}.$$

By Proposition 7.8.3 we see that

$$\mathrm{ex}_q(f) = f((1-q)t, (1-q)qt, (1-q)q^2t, \ldots) \tag{7.28}$$

a minor variant of the stable principal specialization ps. In fact, if $f \in \Lambda^n$ then

$$\mathrm{ex}_q(f) = (1-q)^n t^n \,\mathrm{ps}(f). \tag{7.29}$$

Thus the exponential specialization ex is essentially a limiting case of ps, though we cannot simply set $q = 1$ in (7.29) to conclude that $\mathrm{ex}(f) = 0$ for all f! The substitution $q = 1$ is not a valid formal power series operation, since the operation of setting $q = 1$ in $\mathrm{ps}(f)$ is undefined (or is ∞, if one prefers).

7.9 A Scalar Product

Up to now we have been dealing with a graded algebra Λ with several distinguished bases. We now want to put on Λ the additional structure of a scalar product, i.e., a bilinear form $\Lambda \times \Lambda \to \mathbb{Q}$, which we will denote by $\langle\ ,\ \rangle$. If $\{u_i\}$ and $\{v_j\}$ are bases of a vector space V, then a scalar product on V is uniquely determined by specifying the values $\langle u_i, v_j\rangle$. In particular, we say that $\{u_i\}$ and $\{v_j\}$ are *dual bases* if $\langle u_i, v_j\rangle = \delta_{ij}$ (Kronecker delta) for all i and j. We now define a scalar product on Λ by requiring that $\{m_\lambda\}$ and $\{h_\mu\}$ be dual bases, i.e.,

$$\langle m_\lambda, h_\mu\rangle = \delta_{\lambda\mu}, \tag{7.30}$$

for all $\lambda, \mu \in \mathrm{Par}$. The motivation for this definition will become clear as we develop many desirable and useful properties. First notice that $\langle\ ,\ \rangle$ respects the grading of Λ, in the sense that if f and g are homogeneous then $\langle f, g\rangle = 0$ unless $\deg f = \deg g$.

We now give a series of results which elucidate the nature of the scalar product $\langle\ ,\ \rangle$.

7.9.1 Proposition. *The scalar product $\langle\ ,\ \rangle$ is symmetric, i.e., $\langle f, g\rangle = \langle g, f\rangle$ for all $f, g \in \Lambda$.*

Proof. The result is equivalent to Corollary 7.5.2. More specifically, it suffices by linearity to prove $\langle f, g\rangle = \langle g, f\rangle$ for some bases $\{f\}$ and $\{g\}$ of Λ. Take $\{f\} = \{g\} = \{h_\lambda\}$. Then

$$\langle h_\lambda, h_\mu\rangle = \left\langle \sum_\nu N_{\lambda\nu} m_\nu, h_\mu \right\rangle = N_{\lambda\mu}. \tag{7.31}$$

Since $N_{\lambda\mu} = N_{\mu\lambda}$ by Corollary 7.5.2, we have $\langle h_\lambda, h_\mu\rangle = \langle h_\mu, h_\lambda\rangle$, as desired. □

The following lemma is a basic tool for verifying orthogonality of certain classes of symmetric functions. Its proof is a straightforward exercise in linear algebra and can be omitted without significant loss of understanding.

7.9.2 Lemma. *Let $\{u_\lambda\}$ and $\{v_\lambda\}$ be bases of Λ such that for all $\lambda \vdash n$ we have $u_\lambda, v_\lambda \in \Lambda^n$. Then $\{u_\lambda\}$ and $\{v_\lambda\}$ are dual bases if and only if*

$$\sum_\lambda u_\lambda(x)v_\lambda(y) = \prod_{i,j}(1 - x_i y_j)^{-1}.$$

Proof. Write $m_\lambda = \sum_\rho \zeta_{\lambda\rho} u_\rho$ and $h_\mu = \sum_\nu \eta_{\mu\nu} v_\nu$. Thus

$$\delta_{\lambda\mu} = \langle m_\lambda, h_\mu\rangle = \sum_{\rho,\nu} \zeta_{\lambda\rho}\eta_{\mu\nu}\langle u_\rho, v_\nu\rangle. \tag{7.32}$$

For each fixed $n \geq 0$, regard ζ and η as matrices indexed by Par(n), and let A be the matrix defined by $A_{\rho\nu} = \langle u_\rho, v_\nu\rangle$. Then (7.32) is equivalent to $I = \zeta A \eta^t$, where ^{t} denotes transpose and I the identity matrix. Therefore:

$$\begin{aligned}\{u_\lambda\} \text{ and } \{v_\mu\} \text{ are dual bases} &\iff A = I\\ &\iff I = \zeta\eta^t\\ &\iff I = \zeta^t\eta\\ &\iff \delta_{\rho\nu} = \sum_\lambda \zeta_{\lambda\rho}\eta_{\lambda\nu}.\end{aligned} \tag{7.33}$$

Now by Proposition 7.5.3 we have

$$\begin{aligned}\prod_{i,j}(1 - x_i y_j)^{-1} &= \sum_\lambda m_\lambda(x)h_\lambda(y)\\ &= \sum_\lambda\left(\sum_\rho \zeta_{\lambda\rho}u_\rho(x)\right)\left(\sum_\nu \eta_{\lambda\nu}v_\nu(y)\right)\\ &= \sum_{\rho,\nu}\left(\sum_\lambda \zeta_{\lambda\rho}\eta_{\lambda\nu}\right)u_\rho(x)v_\nu(y).\end{aligned}$$

Since the power series $u_\rho(x)v_\nu(y)$ are linearly independent over $\mathbb{Q}$, the proof follows from (7.33). $\square$

7.9.3 Proposition. *We have*

$$\langle p_\lambda, p_\mu\rangle = z_\lambda\delta_{\lambda\mu}. \tag{7.34}$$

Hence the p_λ's form an orthogonal basis of Λ. (They don't form an orthonormal basis, since $\langle p_\lambda, p_\lambda\rangle \neq 1$.)

Proof. By Proposition 7.7.4 and Lemma 7.9.2 we see that $\{p_\lambda\}$ and $\{p_\mu/z_\mu\}$ are dual bases, which is equivalent to (7.34). $\square$

The length $\|p_\lambda\| = \langle p_\lambda, p_\lambda\rangle^{1/2} = z_\lambda^{1/2}$ is in general not rational. Thus the elements $p_\lambda/\|p_\lambda\|$ form an orthonormal basis of $\Lambda_\mathbb{R}$ but not of Λ (since they don't

belong to Λ). It is natural to ask whether there is a "natural" orthonormal basis for Λ. Even better, is there an *integral* orthonormal basis for Λ, i.e., is there an orthonormal basis $\{b_\lambda\}$ for Λ such that each b_λ is an integer linear combination of m_μ's, and conversely each m_μ is an integer linear combination of b_λ's? Such a basis will thus be a basis for $\Lambda_{\mathbb{Z}}$ (as an abelian group). In Sections 7.10–7.17 we will construct such a basis (see Corollary 7.12.2) and derive many remarkable combinatorial properties that it possesses.

7.9.4 Corollary. *The scalar product $\langle\ ,\ \rangle$ is positive definite, i.e., $\langle f, f\rangle \geq 0$ for all $f \in \Lambda$, with equality if and only if $f = 0$.*

Proof. Write (uniquely) $f = \sum_\lambda c_\lambda p_\lambda$. Then

$$\langle f, f\rangle = \sum c_\lambda^2 z_\lambda.$$

The proof follows since each $z_\lambda > 0$. □

7.9.5 Proposition. *The involution ω is an isometry, i.e., $\langle \omega f, \omega g\rangle = \langle f, g\rangle$ for all $f, g \in \Lambda$.*

Proof. By the bilinearity of the scalar product, it suffices to take $f = p_\lambda$ and $g = p_\mu$. The result then follows from Propositions 7.7.5 and 7.9.3. □

7.10 The Combinatorial Definition of Schur Functions

The four bases m_λ, e_λ, h_λ, and p_λ of Λ discussed in the previous sections all have rather transparent definitions. In this section we consider a fifth basis, whose elements are denoted s_λ and are called *Schur functions*, and whose definition is considerably more subtle. In fact, there are many different (equivalent) ways in which we can define s_λ, viz., in terms of any of the four previous bases, or a "classical" definition involving quotients of determinants, or by abstract properties related to orthogonality and triangularity, or finally by sophisticated algebraic means. All these possible definitions will appear unmotivated to a neophyte. We choose to define s_λ in terms of the m_μ's because this approach is the most combinatorial, though other approaches have their own advantages. In the end, of course, all the approaches produce the same theory.

Much of the importance of Schur functions arises from their connections with such branches of mathematics as representation theory and algebraic geometry. We will discuss the connection with the representation theory of the symmetric group $\mathfrak{S}_n$ in Section 7.18 and with the general linear group $\mathrm{GL}(n, \mathbb{C})$ and related groups in Appendix 2. Another important application of Schur functions not developed here occurs in the Schubert calculus; the cohomology ring of the Grassmann variety $G_k(\mathbb{C}^n)$ can be described in a natural way in terms of Schur functions.

The fundamental combinatorial objects associated with Schur functions are semistandard tableaux. Let λ be a partition. A *semistandard (Young) tableau* (SSYT) of *shape* λ is an array $T = (T_{ij})$ of positive integers of shape λ (i.e., $1 \le i \le \ell(\lambda)$, $1 \le j \le \lambda_i$) that is weakly increasing in every row and strictly increasing in every column. The *size* of an SSYT is its number of entries. An example of an SSYT of shape $(6,5,3,3)$ is given by

$$\begin{array}{llllll} 1&1&1&3&4&4\\ 2&4&4&5&5&\\ 5&5&7&&&\\ 6&9&9.&&& \end{array}$$

If T is an SSYT of shape λ then we write $\lambda = \mathrm{sh}(T)$. Hence the size of T is just $|\mathrm{sh}(T)|$. We may also think of an SSYT of shape λ as the Young diagram (as defined in Section 1.3) of λ whose boxes have been filled with positive integers (satisfying certain conditions). For instance, the above SSYT may be written

1	1	1	3	4	4
2	4	4	5	5	
5	5	7			
6	9	9			

We say that T has *type* $\alpha = (\alpha_1, \alpha_2, \ldots)$, denoted $\alpha = \mathrm{type}(T)$, if T has $\alpha_i = \alpha_i(T)$ parts equal to i. Thus the above SSYT has type $(3,1,1,4,4,1,1,0,2)$. For any SSYT T of type α (or indeed for any multiset on $\mathbb{P}$ with possible additional structure), write

$$x^T = x_1^{\alpha_1(T)} x_2^{\alpha_2(T)} \cdots.$$

For our running example we have

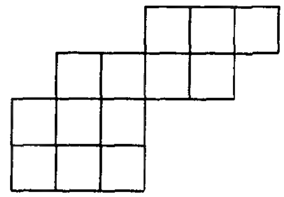
$$x^T = x_1^3 x_2 x_3 x_4^4 x_5^4 x_6 x_7 x_9^2.$$

There is a generalization of SSYTs of shape λ that fits naturally into the theory of symmetric functions. If λ and μ are partitions with $\mu \subseteq \lambda$ (i.e., $\mu_i \le \lambda_i$ for all i), then define a *semistandard tableau* of *(skew) shape* λ/μ to be an array $T = (T_{ij})$ of positive integers of shape λ/μ (i.e., $1 \le i \le \ell(\lambda)$, $\mu_i < j \le \lambda_i$) that is weakly increasing in every row and strictly increasing in every column. An example of an SSYT of shape $(6,5,3,3)/(3,1)$ is given by

$$\begin{array}{llllll} &&&3&4&4\\ &1&4&7&7&\\ 2&2&6&&&\\ 3&8&8.&&& \end{array}$$

We can similarly extend the definition of a Young diagram of shape λ to one of

shape λ/μ. Thus the Young diagram of shape $(6, 5, 3, 3)/(3, 1)$ is given by



Thus an SSYT of shape λ/μ may be regarded as a Young diagram of shape λ/μ whose boxes have been filled with positive integers (satisfying certain conditions), just as for "ordinary shapes" λ. For instance, the above SSYT of shape $(6, 5, 3, 3)/(3, 1)$ may be written

$$\begin{array}{cccccc} & & & 3 & 4 & 4 \\ & 1 & 4 & 7 & 7 & \\ 2 & 2 & 6 & & & \\ 3 & 8 & 8 & & & \end{array}$$

The definitions of type(T) and x^T carry over directly from SSYTs T of ordinary shape to those of skew shape.

We now come to the key definition of this entire chapter. As mentioned previously, this definition will appear entirely unmotivated until we proceed further.

7.10.1 Definition. Let λ/μ be a skew shape. The *skew Schur function* $s_{\lambda/\mu} = s_{\lambda/\mu}(x)$ of *shape* λ/μ in the variables $x = (x_1, x_2, \ldots)$ is the formal power series

$$s_{\lambda/\mu}(x) = \sum_T x^T,$$

summed over all SSYTs T of shape λ/μ. If $\mu = \emptyset$ so $\lambda/\mu = \lambda$, then we call $s_\lambda(x)$ the *Schur function* of shape λ.

For instance, the SSYTs of shape $(2, 1)$ with largest part at most three are given by

$$\begin{array}{cccccccc} 11 & 12 & 11 & 13 & 22 & 23 & 12 & 13 \\ 2 & 2 & 3 & 3 & 3 & 3 & 3 & 2 \end{array}$$

Hence

$$\begin{aligned} s_{21}(x_1, x_2, x_3) &= x_1^2x_2 + x_1x_2^2 + x_1^2x_3 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2 + 2x_1x_2x_3 \\ &= m_{21}(x_1, x_2, x_3) + 2m_{111}(x_1, x_2, x_3). \end{aligned}$$

Thus, since at most three distinct variables can occur in a term of s_{21}, we have $s_{21} = m_{21} + 2m_{111}$ (as elements of Λ, i.e., as symmetric functions in *infinitely* many variables). It is by no means obvious that $s_{\lambda/\mu}$ is in fact always a symmetric function.

7.10.2 Theorem. *For any skew shape λ/μ, the skew Schur function $s_{\lambda/\mu}$ is a symmetric function.*

Proof. It suffices to show [why?] that $s_{\lambda/\mu}$ is invariant under interchanging x_i and x_{i+1}. Suppose that $|\lambda/\mu| = n$ and that $\alpha = (\alpha_1, \alpha_2, \ldots)$ is a weak composition of n. Let

$$\tilde{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{i-1}, \alpha_{i+1}, \alpha_i, \alpha_{i+2}, \ldots).$$

If $\mathcal{T}_{\lambda/\mu,\alpha}$ denotes the set of all SSYTs of shape λ/μ and type α, then we seek a bijection $\varphi : \mathcal{T}_{\lambda/\mu,\alpha} \to \mathcal{T}_{\lambda/\mu,\tilde{\alpha}}$.

Let $T \in \mathcal{T}_{\lambda/\mu,\alpha}$. Consider the parts of T equal to i or $i+1$. Some columns of T will contain no such parts, while some others will contain two such parts, viz., one i and one $i+1$. These columns we ignore. The remaining parts equal to i or $i+1$ occur once in each column, and consist of rows with a certain number r of i's followed by a certain number s of $i+1$'s. (Of course r and s depend on the row in question.) For example, a portion of T could look as follows:

$$\begin{array}{ccccc} & & & & i\\ i & i & \underbrace{i\ \ i}_{r=2} & \underbrace{i+1\ \ i+1\ \ i+1\ \ i+1}_{s=4} & i+1\\ i+1 & i+1 & & & \end{array}$$

In each such row convert the r i's and s $i+1$'s to s i's and r $i+1$'s:

$$\begin{array}{ccccc} & & & & i\\ i & i & \underbrace{i\ \ i\ \ i\ \ i}_{s=4} & \underbrace{i+1\ \ i+1}_{r=2} & i+1\\ i+1 & i+1 & & & \end{array}$$

It's easy to see that the resulting array $\varphi(T)$ belongs to $\mathcal{T}_{\lambda/\mu,\tilde{\alpha}}$, and that φ establishes the desired bijection. □

If $\lambda \vdash n$ and α is a weak composition of n, then let $K_{\lambda\alpha}$ denote the number of SSYTs of shape λ and type α. $K_{\lambda\alpha}$ is called a *Kostka number* and plays a prominent role in the theory of symmetric functions. By Definition 7.10.1 we have

$$s_\lambda = \sum_\alpha K_{\lambda\alpha} x^\alpha,$$

summed over all weak compositions α of n, so by Theorem 7.10.2 we have

$$s_\lambda = \sum_{\mu \vdash n} K_{\lambda\mu} m_\mu. \tag{7.35}$$

More generally, we can define the *skew Kostka number* $K_{\lambda/\nu,\alpha}$ as the number of SSYTs of shape λ/ν and type α, so that if $|\lambda/\nu| = n$ then

$$s_{\lambda/\nu} = \sum_{\mu \vdash n} K_{\lambda/\nu,\mu} m_\mu. \tag{7.36}$$

No simple formula is known in general for $K_{\lambda/\nu,\mu}$, or even $K_{\lambda\mu}$, and it is unlikely that such a formula exists. For certain λ, ν, and μ a formula can be given, the most

important being the case $\nu = \emptyset$ and $\mu = \langle 1^n \rangle$. While we will give this formula later (Corollary 7.21.6), let us here consider more closely the combinatorial significance of the number $K_{\lambda,1^n}$, also denoted f^λ. By definition, f^λ is the number of ways to insert the numbers $1, 2, \ldots, n$ into the shape $\lambda \vdash n$, each number appearing once, so that every row and column is increasing. Such an array is called a *standard Young tableau* (SYT) (or just *standard tableau*) of shape λ. For instance, the SYTs of shape $(3, 2)$ are

$$\begin{matrix} 123 & 124 & 125 & 134 & 135 \\ 45 & 35 & 34 & 25 & 24 \end{matrix},$$

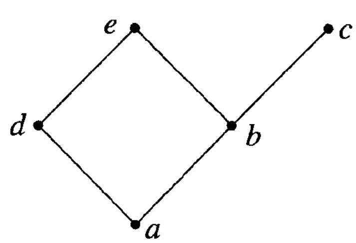
so $f^{(3,2)} = 5$. The number f^λ has several alternative combinatorial interpretations, as given by the following proposition.

7.10.3 Proposition. *Let $\lambda \in$ Par. Then the number f^λ counts the objects in items (a)–(e) below. We illustrate these objects with the case $\lambda = (3, 2)$.*

(a) Chains of partitions. *Saturated chains in the interval $[\emptyset, \lambda]$ of Young's lattice Y, or equivalently, sequences $\emptyset = \lambda^0, \lambda^1, \ldots, \lambda^n = \lambda$ of partitions (which we identify with their diagrams) such that λ^i is obtained from λ^{i-1} by adding a single square.*

$$\begin{array}{l} \emptyset \subset 1 \subset 2 \subset 3 \subset 31 \subset 32 \\ \emptyset \subset 1 \subset 2 \subset 21 \subset 31 \subset 32 \\ \emptyset \subset 1 \subset 2 \subset 21 \subset 22 \subset 32 \\ \emptyset \subset 1 \subset 11 \subset 21 \subset 31 \subset 32 \\ \emptyset \subset 1 \subset 11 \subset 21 \subset 22 \subset 32 \end{array}$$

(b) Linear extensions. *Let P_λ be the poset whose elements are the squares of the diagram of λ, with t covering s if t lies directly to the right or directly below s (with no squares in between). Such posets are just the finite order ideals of $\mathbb{N} \times \mathbb{N}$. Then $f^\lambda = e(P_\lambda)$, the number of linear extensions of P_λ.*



abcde
abdce
abdec
adbce
adbec

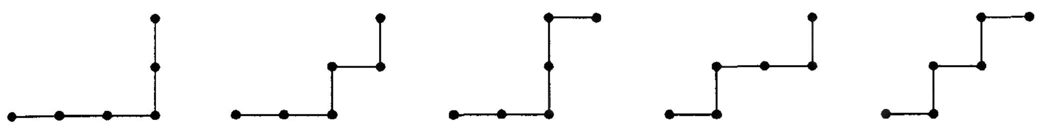
(c) Ballot sequences. *Ways in which n voters can vote sequentially in an election for candidates $A_1, A_2, \ldots$, so that for all i, A_i receives λ_i votes, and so that A_i never trails A_{i+1} in the voting. (We denote such a voting sequence as $a_1 a_2 \cdots a_n$, where the k-th voter votes for A_{a_k}.)*

11122 11212 11221 12112 12121

(d) Lattice permutations. *Sequences $a_1 a_2 \cdots a_n$ in which i occurs λ_i times, and such that in any left factor $a_1 a_2 \cdots a_j$, the number of i's is at least as great as the number of $i+1$'s (for all i). Such a sequence is called a* lattice permutation *(or* Yamanouchi word *or* ballot sequence*) of type λ.*

11122 11212 11221 12112 12121

(e) Lattice paths. *Lattice paths $0 = v_0, v_1, \ldots, v_n$ in $\mathbb{R}^\ell$ (where $\ell = \ell(\lambda)$) from the origin v_0 to $v_n = (\lambda_1, \lambda_2, \ldots, \lambda_\ell)$, with each step a unit coordinate vector, and staying within the region (or cone) $x_1 \geq x_2 \geq \cdots \geq x_\ell \geq 0$.*



Proof. (a) Insert i into the square that was added to λ^{i-1} in order to obtain λ^i, to get an SYT of shape λ.
(b) The interval $[\emptyset, \lambda]$ in Y is just $J(P_\lambda)$, the lattice of order ideals of P_λ, so the equivalence between our interpretations (a) and (b) of f^λ is just a special case of the discussion following Proposition 3.5.2.
(c) If the k-th voter votes for A_i, then put k in the i-th row of the shape λ.
(d) Clearly the voting sequences in (c) are identical to the lattice permutations of (d).
(e) If $a_1 a_2 \cdots a_n$ is a lattice permutation as in (d), then let $v_i - v_{i-1}$ be the a_i-th unit coordinate vector (i.e., the vector with a one in position a_i and zeros elsewhere) to obtain a lattice path. Alternatively, the equivalence between (b) and (e) is a special case of the discussion preceding Example 3.5.3. □

All five of the above interpretations can be straightforwardly generalized to the skew case $f^{\lambda/\mu}$. We leave the details of this task to the interested reader.

There is a combinatorial object equivalent to an SSYT that is worth mentioning. A *Gelfand–Tsetlin pattern* (sometimes called just a *Gelfand pattern*), or *complete branching*, is a triangular array G of nonnegative integers, say

$$\begin{array}{ccccccccc}
a_{11} & & a_{12} & & a_{13} & & \cdots & & a_{1n} \\
& a_{22} & & a_{23} & & \cdots & & a_{2n} & \\
& & a_{33} & & \cdots & & a_{3n} & & \\
& & & \ddots & & \iddots & & & \\
& & & & a_{nn} & & & &
\end{array} \tag{7.37}$$

such that $a_{ij} \leq a_{i+1,j+1} \leq a_{i,j+1}$ when all three numbers are defined. In other words, the rows of G are weakly increasing, and $a_{i+1,j+1}$ lies weakly between its

two neighbors above it. An example of a Gelfand–Tsetlin pattern is

$$\begin{array}{ccccccccc} 0 & & 2 & & 2 & & 3 & & 6 \\ & 0 & & 2 & & 2 & & 4 & \\ & & 1 & & 2 & & 4 & & \\ & & & 1 & & 3 & & & \\ & & & & 3. & & & & \end{array}$$

Given the Gelfand–Tsetlin pattern G of equation (7.37), let λ^i be the i-th row of G in reverse order. Define a tableau $T = T(G)$ by inserting $n - i + 1$ into the squares of the skew shape λ^i/λ^{i+1}. For the example above, $T(G)$ is given by

$$\begin{array}{llllll} 1 & 1 & 1 & 3 & 5 & 5 \\ 2 & 3 & 5 & & & \\ 3 & 4 & & & & \\ 5 & 5. & & & & \end{array}$$

We obtain an SSYT of shape λ^1 (the first row of G in reverse order) and largest part at most n. This correspondence between Gelfand–Tsetlin patterns with fixed first row α of length n and SSYT of shape α^r (the elements of α in reverse order) and largest part at most n is easily seen to be a bijection.

It is sometimes more convenient in dealing with Schur functions, Kostka numbers, etc., to work with arrays that are *decreasing* in rows and columns rather than with SSYT. Define a *reverse SSYT* or *column-strict plane partition* (sometimes abbreviated as *costripp*) of (skew) shape λ/μ to be an array of positive integers of shape λ/μ that is weakly decreasing in rows and strictly decreasing in columns. Define the *type* α of a reverse SSYT exactly as for ordinary SSYT. For instance, the array

$$\begin{array}{cccccc} & & & 6 & 5 & 5 \\ & 8 & 5 & 2 & 2 & \\ 7 & 7 & 3 & & & \\ 6 & 1 & 1 & & & \end{array}$$

is a reverse SSYT of shape $(6, 5, 3, 3)/(3, 1)$ and type $(2, 2, 1, 0, 3, 2, 2, 1)$.

Define $\hat{K}_{\lambda/\mu,\alpha}$ to be the number of reverse SSYTs of shape λ/μ and type α. The next proposition shows that for many purposes there is no significant difference between ordinary and reverse SSYTs.

7.10.4 Proposition. *Let λ/μ be a skew partition of n, and let α be a weak composition of n. Then $\hat{K}_{\lambda/\mu,\alpha} = K_{\lambda/\mu,\alpha}$.*

Proof. Suppose that T is a reverse SSYT of shape λ and type $\alpha = (\alpha_1, \alpha_2, \ldots)$. Let k denote the largest part of T. The transformation $T_{ij} \mapsto k + 1 - T_{ij}$ shows

that $\hat{K}_{\lambda\alpha} = K_{\lambda\bar{\alpha}}$, where $\bar{\alpha} = (\alpha_k, \alpha_{k-1}, \ldots, \alpha_1, 0, 0, \ldots)$. But by Theorem 7.10.2 we have $K_{\lambda\bar{\alpha}} = K_{\lambda\alpha}$, and the proof is complete. □

We now wish to show that the Schur functions s_λ form a $\mathbb{Q}$-basis for Λ. The following proposition implies an even stronger result.

7.10.5 Proposition. *Suppose that μ and λ are partitions with $|\mu| = |\lambda|$ and $K_{\lambda\mu} \neq 0$. Then $\mu \leq \lambda$ (dominance order). Moreover, $K_{\lambda\lambda} = 1$.*

Proof. Suppose that $K_{\lambda\mu} \neq 0$. By definition, there exists an SSYT T of shape λ and type μ. Suppose that a part $T_{ij} = k$ appears below the k-th row (i.e., $i > k$). Then we have $1 \leq T_{1k} < T_{2k} < \cdots < T_{ik} = k$ for $i > k$, which is impossible. Hence the parts $1, 2, \ldots, k$ all appear in the first k rows, so $\mu_1 + \mu_2 + \cdots + \mu_k \leq \lambda_1 + \lambda_2 + \cdots + \lambda_k$, as desired. Moreover, if $\mu = \lambda$ then we must have $T_{ij} = i$ for all (i, j), so $K_{\lambda\lambda} = 1$. □

7.10.6 Corollary. *The Schur functions s_λ with $\lambda \in \mathrm{Par}(n)$ form a basis for Λ^n, so $\{s_\lambda : \lambda \in \mathrm{Par}\}$ is a basis for Λ. In fact, the transition matrix $(K_{\lambda\mu})$ which expresses the s_λ's in terms of the m_μ's, with respect to any linear ordering of $\mathrm{Par}(n)$ that extends dominance order, is lower triangular with 1's on the main diagonal.*

$$
\begin{aligned}
s_1 &= m_1 \\[1ex]
s_{11} &= m_{11} \\
s_2 &= m_{11} + m_2 \\[1ex]
s_{111} &= m_{111} \\
s_{21} &= 2m_{111} + m_{21} \\
s_3 &= m_{111} + m_{21} + m_3 \\[1ex]
s_{1111} &= m_{1111} \\
s_{211} &= 3m_{1111} + m_{211} \\
s_{22} &= 2m_{1111} + m_{211} + m_{22} \\
s_{31} &= 3m_{1111} + 2m_{211} + m_{22} + m_{31} \\
s_4 &= m_{1111} + m_{211} + m_{22} + m_{31} + m_4 \\[1ex]
s_{11111} &= m_{11111} \\
s_{2111} &= 4m_{11111} + m_{2111} \\
s_{221} &= 5m_{11111} + 2m_{2111} + m_{221} \\
s_{311} &= 6m_{11111} + 3m_{2111} + m_{221} + m_{311} \\
s_{32} &= 5m_{11111} + 3m_{2111} + 2m_{221} + m_{311} + m_{32} \\
s_{41} &= 4m_{11111} + 3m_{2111} + 2m_{221} + 2m_{311} + m_{32} + m_{41} \\
s_5 &= m_{11111} + m_{2111} + m_{221} + m_{311} + m_{32} + m_{41} + m_5
\end{aligned}
$$

Figure 7-4. The Kostka numbers $K_{\lambda\mu}$.

Proof. Proposition 7.10.5 is equivalent to the assertion about $(K_{\lambda\mu})$. Since a lower triangular matrix with 1's on the main diagonal is invertible, it follows that $\{s_\lambda : \lambda \in \mathrm{Par}(n)\}$ is a $\mathbb{Q}$-basis for Λ^n. □

Note that in fact $\{s_\lambda : \lambda \in \mathrm{Par}(n)\}$ is a $\mathbb{Z}$-basis for $\Lambda^n_{\mathbb{Z}}$, since each $K_{\lambda\lambda} = 1$, rather than just $K_{\lambda\lambda} \neq 0$.

In subsequent sections we will work out the basic theory of Schur functions, that is, the transition matrices between the s_λ's and the bases $m_\lambda, h_\lambda, e_\lambda, p_\lambda$, as well as connections with the scalar product $\langle\ ,\ \rangle$ and the automorphism ω. (We have already considered, essentially by definition, the transition matrix $(K_{\lambda\mu})$ from the m_λ's to the s_λ's, but we don't know yet what the inverse matrix looks like.) We will also give several enumerative applications of the theory of symmetric functions: the enumeration of plane partitions, some results on permutation statistics, and Pólya's theory of enumeration under group action.

Figure 7-4 gives a short table of the Kostka numbers $K_{\lambda\mu}$.

7.11 The RSK Algorithm

There is a remarkable combinatorial correspondence associated with the theory of symmetric functions, called the *RSK algorithm*. (For the meaning of the initials RSK, as well as for other names of the algorithm, see the Notes at the end of this chapter.) We will develop here only the most essential properties of the RSK algorithm, thereby allowing us to give combinatorial proofs of some fundamental properties of Schur functions. It is also possible to give purely algebraic proofs of these results, but of course in a text on enumerative combinatorics we prefer combinatorial proofs.

The basic operation of the RSK algorithm consists of the *row insertion* $P \leftarrow k$ of a positive integer k into a nonskew SSYT $P = (P_{ij})$. The operation $P \leftarrow k$ is defined as follows: Let r be the largest integer such that $P_{1,r-1} \leq k$. (If $P_{11} > k$ then let $r = 1$.) If P_{1r} doesn't exist (i.e., P has $r-1$ columns), then simply place k at the end of the first row. The insertion process stops, and the resulting SSYT is $P \leftarrow k$. If, on the other hand, P has at least r columns, so that P_{1r} exists, then replace P_{1r} by k. The element then "bumps" $P_{1r} := k'$ into the second row, i.e., insert k' into the second row of P by the insertion rule just described. Continue until an element is inserted at the end of a row (possibly as the first element of a new row). The resulting array is $P \leftarrow k$.

7.11.1 Example. Let

$$P = \begin{matrix} 1\,1\,2\,4\,5\,5\,6 \\ 2\,3\,3\,6\,6\,8 \\ 4\,4\,6\,8 \\ 6\,7 \\ 8\,9. \end{matrix}$$

Then $P \leftarrow 4$ is shown below, with the elements inserted into each row (either by bumping or by the final insertion in the fourth row) in boldface. Thus the 4 bumps the 5, the 5 bumps the 6, the 6 bumps the 8, and the 8 is inserted at the end of a row. The set of positions of these boldface elements is called the *insertion path* $I(P \leftarrow 4)$ of 4 (the number being inserted into P). Thus for this example we have $I(P \leftarrow 4) = \{(1, 5), (2, 4), (3, 4), (4, 3)\}$:

$$\begin{array}{l} 1\,1\,2\,4\,\mathbf{4}\,5\,6 \\ 2\,3\,3\,\mathbf{5}\,6\,8 \\ 4\,4\,6\,\mathbf{6} \\ 6\,7\,\mathbf{8} \\ 8\,9. \end{array}$$

There are two technical properties of insertion paths that are of great use in proving properties of the RSK algorithm.

7.11.2 Lemma. (a) *When we insert k into an SSYT P, then the insertion path moves to the left. More precisely, if $(r, s), (r + 1, t) \in I(P \leftarrow k)$ then $t \leq s$.*
(b) *Let P be an SSYT, and let $j \leq k$. Then $I(P \leftarrow j)$ lies strictly to the left of $I((P \leftarrow j) \leftarrow k)$. More precisely, if $(r, s) \in I(P \leftarrow j)$ and $(r, t) \in I((P \leftarrow j) \leftarrow k)$, then $s < t$. Moreover, $I((P \leftarrow j) \leftarrow k)$ does not extend below the bottom of $I(P \leftarrow j)$. Equivalently,*

$$\#I((P \leftarrow j) \leftarrow k) \leq \#I(P \leftarrow j).$$

Proof. (a) Suppose that $(r, s) \in I(P \leftarrow k)$. Now either $P_{r+1,s} > P_{rs}$ (since P is strictly increasing in columns) or else there is no $(r + 1, s)$ entry of P. In the first case, P_{rs} cannot get bumped to the right of column s without violating the fact that the rows of $P \leftarrow k$ are weakly increasing, since P_{rs} would be to the right of $P_{r+1,s}$ on the same row. The second case is clearly impossible, since we would otherwise have a gap in row $r + 1$. Hence (a) is proved.
(b) Since a number can only bump a strictly larger number, it follows that k is inserted in the first row of $P \leftarrow j$ strictly to the right of j. Since the first row of P is weakly increasing, j bumps an element no larger than the element k bumps. Hence by induction $I(P \leftarrow j)$ lies strictly to the left of $I((P \leftarrow j) \leftarrow k)$. The bottom element b of $I(P \leftarrow j)$ was inserted at the end of its row. By what was just proved, if $I((P \leftarrow j) \leftarrow k)$ has an element c in this row, then it lies to the right of b. Hence c was inserted at the end of the row, so the insertion procedure terminates. It follows that $I((P \leftarrow j) \leftarrow k)$ can never go below the bottom of $I(P \leftarrow j)$. □

7.11.3 Corollary. *If P is an SSYT and $k \geq 1$, then $P \leftarrow k$ is also an SSYT.*

Proof. It is clear that the rows of $P \leftarrow k$ are weakly increasing. Now a number a can only bump a larger number b. By Lemma 7.11.2(a), b does not move to the right when it is bumped. Hence b is inserted below a number that is strictly smaller than b, so $P \leftarrow k$ remains an SSYT. □

Now let $A = (a_{ij})_{i,j\geq 1}$ be an $\mathbb{N}$-matrix with finitely many nonzero entries. We will say that A is an $\mathbb{N}$-matrix of *finite support*. We can think of A as either an infinite matrix or as an $m \times n$ matrix when $a_{ij} = 0$ for $i > m$ and $j > n$. Associate with A a *generalized permutation* or *two-line array* w_A defined by

$$w_A = \begin{pmatrix} i_1 & i_2 & i_3 & \cdots & i_m \\ j_1 & j_2 & j_3 & \cdots & j_m \end{pmatrix}, \tag{7.38}$$

where (a) $i_1 \leq i_2 \leq \cdots \leq i_m$, (b) if $i_r = i_s$ and $r \leq s$, then $j_r \leq j_s$, and (c) for each pair (i, j), there are exactly a_{ij} values of r for which $(i_r, j_r) = (i, j)$. It is easily seen that A determines a unique two-line array w_A satisfying (a)–(c), and conversely any such array corresponds to a unique A. For instance, if

$$A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix}, \tag{7.39}$$

then the corresponding two-line array is

$$w_A = \begin{pmatrix} 1\,1\,1\,2\,2\,3\,3 \\ 1\,3\,3\,2\,2\,1\,2 \end{pmatrix}. \tag{7.40}$$

We now associate with A (or w_A) a pair (P, Q) of SSYTs of the same shape, as follows. Let w_A be given by (7.38). Begin with $(P(0), Q(0)) = (\emptyset, \emptyset)$ (where $\emptyset$ denotes the empty SSYT). If $t < m$ and $(P(t), Q(t))$ are defined, then let

(a) $P(t+1) = P(t) \leftarrow j_{t+1}$;
(b) $Q(t+1)$ be obtained from $Q(t)$ by inserting i_{t+1} (leaving all parts of $Q(t)$ unchanged) so that $P(t+1)$ and $Q(t+1)$ have the same shape.

The process ends at $(P(m), Q(m))$, and we define $(P, Q) = (P(m), Q(m))$. We denote this correspondence by $A \xrightarrow{\text{RSK}} (P, Q)$ and call it the *RSK algorithm*. We call P the *insertion tableau* and Q the *recording tableau* of A or of w_A.

7.11.4 Example. Let A and w_A be given by (7.39) and (7.40). The SSYTs $(P(1), Q(1)), \dots, (P(7), Q(7)) = (P, Q)$ are as follows:

$P(i)$	$Q(i)$
1	1
1 3	1 1
1 3 3	1 1 1
1 2 3 3	1 1 1 2
1 2 2 3 3	1 1 1 2 2
1 1 2 2 3 3	1 1 1 2 2 3
1 1 2 2 2 3 3	1 1 1 3 2 2 3.

The main result on the RSK algorithm is the following.

7.11.5 Theorem. *The RSK algorithm is a bijection between $\mathbb{N}$-matrices $A = (a_{ij})_{i,j\geq 1}$ of finite support and ordered pairs (P, Q) of SSYT of the same shape. In this correspondence,*

$$j \text{ occurs in } P \text{ exactly } \sum_i a_{ij} \text{ times} \tag{7.41}$$

$$i \text{ occurs in } Q \text{ exactly } \sum_j a_{ij} \text{ times.} \tag{7.42}$$

(These last two conditions are equivalent to $\operatorname{type}(P) = \operatorname{col}(A)$, $\operatorname{type}(Q) = \operatorname{row}(A)$.*)*

Proof. By Corollary 7.11.3, P is an SSYT. Clearly, by definition of the RSK algorithm P and Q have the same shape, and also (7.41) and (7.42) hold. Thus we must show the following: (a) Q is an SSYT, and (b) the RSK algorithm is a bijection, i.e., given (P, Q), one can uniquely recover A.

To prove (a), first note that since the elements of Q are inserted in weakly increasing order, it follows that the rows and columns of Q are weakly increasing. Thus we must show that the columns of Q are strictly increasing, i.e., no two equal elements of the top row of w_A can end up in the same column of Q. But if $i_k = i_{k+1}$ in the top row, then we must have $j_k \leq j_{k+1}$. Hence by Lemma 7.11.2(b),

the insertion path of j_{k+1} will always lie strictly to the right of the path for j_k, and will never extend below the bottom of j_k's insertion path. It follows that the bottom elements of the two insertion paths lie in different columns, so the columns of Q are strictly increasing as desired.

The above argument establishes an important property of the RSK algorithm: *Equal elements of Q are inserted strictly left to right.*

It remains to show that the RSK algorithm is a bijection. Thus given $(P, Q) = (P(m), Q(m))$, let Q_{rs} be the rightmost occurrence of the largest entry of Q (where Q_{rs} is the element of Q in row r and column s). Since equal elements of Q are inserted left to right, it follows that $Q_{rs} = i_m$, $Q(m-1) = Q(m)\backslash Q_{rs}$ (i.e., $Q(m)$ with the element Q_{rs} deleted), and that P_{rs} was the last element of P to be bumped into place after inserting j_m into $P(m-1)$. But it is then easy to reverse the insertion procedure $P(m-1) \leftarrow j_m$. P_{rs} must have been bumped by the rightmost element $P_{r-1,t}$ of row $r-1$ of P that is smaller than P_{rs}. Hence remove P_{rs} from P, replace $P_{r-1,t}$ with P_{rs}, and continue by replacing the rightmost element of row $r-2$ of P that is smaller than $P_{r-1,t}$ with $P_{r-1,t}$, etc. Eventually some element j_m is removed from the first row of P. We have thus uniquely recovered (i_m, j_m) and $(P(m-1), Q(m-1))$. By iterating this procedure we recover the entire two-line array w_A. Hence the RSK algorithm is injective.

To show surjectivity, we need to show that applying the procedure of the previous paragraph to an arbitrary pair (P, Q) of SSYTs of the same shape always yields a valid two-line array

$$w_A = \begin{pmatrix} i_1 & \cdots & i_m \\ j_1 & \cdots & j_m \end{pmatrix}.$$

Clearly $i_1 \leq i_2 \leq \cdots \leq i_m$, so we need to show that if $i_k = i_{k+1}$ then $j_k \leq j_{k+1}$. Let $i_k = Q_{rs}$ and $i_{k+1} = Q_{uv}$, so $r \geq u$ and $s < v$. When we begin to apply inverse bumping to P_{uv}, it occupies the end of its row (row u). Hence when we apply inverse bumping to P_{rs}, its "inverse insertion path" intersects row u strictly to the left of column v. Thus at row u the inverse insertion path of P_{rs} lies strictly to the left of that of P_{uv}. By a simple induction argument (essentially the "inverse" of Lemma 7.11.2(b)), the entire inverse insertion path of P_{rs} lies strictly to the left of that of P_{uv}. In particular, before removing i_{k+1} the two elements j_k and j_{k+1} appear in the first row with j_k to the left of j_{k+1}. Hence $j_k \leq j_{k+1}$ as desired, completing the proof. □

In Section 7.13 we will give an alternative "geometric" description of the RSK algorithm useful in proving some remarkable properties. This geometric description is only defined when the matrix A is a *permutation matrix*, i.e., an $n \times n$ $(0, 1)$-matrix with exactly one 1 in every row and column. In this case the top line of the two-line array is just $1\,2 \cdots n$, while the bottom line is a permutation w of $1, 2, \ldots, n$ that we can identify with A. When the RSK algorithm is applied to a

permutation matrix A (or permutation $w \in \mathfrak{S}_n$), the resulting tableaux P, Q are just standard Young tableaux (of the same shape). Conversely, if P and Q are SYTs of the same shape, then the matrix A satisfying $A \xrightarrow{\text{RSK}} (P, Q)$ is a permutation matrix. Hence the RSK algorithm sets up a bijection between the symmetric group $\mathfrak{S}_n$ and pairs (P, Q) of SYTs of the same shape $\lambda \vdash n$. In particular, if f^λ denotes the number of SYTs of shape λ, then we have the fundamental identity

$$\sum_{\lambda \vdash n} (f^\lambda)^2 = n!. \tag{7.43}$$

Although permutation matrices are very special cases of $\mathbb{N}$-matrices of finite support, in fact the RSK algorithm for arbitrary $\mathbb{N}$-matrices A can be reduced to the case of permutation matrices. Namely, given the two-line array w_A, say of length n, replace the first row by $1, 2, \ldots, n$. Suppose that the second row of w_A has c_i i's. Then replace the 1's in the second row from left-to-right with $1, 2, \ldots, c_1$, next the 2's from left-to-right with $c_1 + 1, c_1 + 2, \ldots, c_1 + c_2$, etc., until the second row becomes a permutation of $1, 2, \ldots, n$. Denote the resulting two-line array by $\tilde{w}_A$. For instance, if

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 3 & 0 \end{bmatrix},$$

then

$$w_A = \begin{pmatrix} 1 & 1 & 1 & 2 & 2 & 3 & 3 & 3 & 3 \\ 1 & 1 & 3 & 2 & 3 & 1 & 2 & 2 & 2 \end{pmatrix},$$

and w_A is replaced by

$$\tilde{w}_A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 8 & 4 & 9 & 3 & 5 & 6 & 7 \end{pmatrix}.$$

7.11.6 Lemma. *Let*

$$w_A = \begin{pmatrix} i_1 & i_2 & \cdots & i_n \\ j_1 & j_2 & \cdots & j_n \end{pmatrix}$$

be a two-line array, and let

$$\tilde{w}_A = \begin{pmatrix} 1 & 2 & \cdots & n \\ \tilde{j}_1 & \tilde{j}_2 & \cdots & \tilde{j}_n \end{pmatrix}.$$

Suppose that $\tilde{w}_A \xrightarrow{\text{RSK}} (\tilde{P}, \tilde{Q})$. *Let* (P, Q) *be the tableaux obtained from* $\tilde{P}$ *and* $\tilde{Q}$ *by replacing* k *in* $\tilde{Q}$ *by* i_k, *and* $\tilde{j}_k$ *in* $\tilde{P}$ *by* j_k. *Then* $w_A \xrightarrow{\text{RSK}} (P, Q)$. *In other words, the operation* $w_A \mapsto \tilde{w}_A$ *"commutes" with the RSK algorithm.*

Proof. Suppose that when the number j is inserted into a row at some stage of the RSK algorithm, it occupies the k-th position in the row. If this number j were replaced by a larger number $j + \epsilon$, smaller than any element of the row which is greater than j, then $j + \epsilon$ would also be inserted at the k-th position. From this we see that the insertion procedure for elements $j_1 j_2 \cdots j_n$ exactly mimics that for $\tilde{j}_1 \tilde{j}_2 \cdots \tilde{j}_n$, and the proof follows. $\square$

The process of replacing w_A with $\tilde{w}_A$, P with $\tilde{P}$, etc., is called *standardization.* Compare the second proof of Proposition 1.3.17.

7.12 Some Consequences of the RSK Algorithm

The most important result concerning symmetric functions that follows directly from the RSK algorithm is the following, known as the *Cauchy identity.*

7.12.1 Theorem. *We have*

$$\prod_{i,j}(1 - x_i y_j)^{-1} = \sum_{\lambda} s_\lambda(x) s_\lambda(y). \tag{7.44}$$

Proof. Write

$$\prod_{i,j}(1 - x_i y_j)^{-1} = \prod_{i,j}\left[\sum_{a_{ij}\geq 0}(x_i y_j)^{a_{ij}}\right]. \tag{7.45}$$

A term $x^\alpha y^\beta$ in this expansion is obtained by choosing an $\mathbb{N}$-matrix $A^t = (a_{ij})^t$ (the transpose of A) of finite support with $\text{row}(A) = \alpha$ and $\text{col}(A) = \beta$. Hence the coefficient of $x^\alpha y^\beta$ in (7.45) is the number $N_{\alpha\beta}$ of $\mathbb{N}$-matrices A with $\text{row}(A) = \alpha$ and $\text{col}(A) = \beta$. (This statement is also equivalent to (7.9).) On the other hand, the coefficient of $x^\alpha y^\beta$ in $\sum_\lambda s_\lambda(x) s_\lambda(y)$ is the number of pairs (P, Q) of SSYT of the same shape λ such that $\text{type}(P) = \alpha$ and $\text{type}(Q) = \beta$. The RSK algorithm sets up a bijection between the matrices A and the tableau pairs (P, Q), so the proof follows. $\square$

The Cauchy identity (7.44) has a number of immediate corollaries.

7.12.2 Corollary. *The Schur functions form an orthonormal basis for* Λ, *i.e.,* $\langle s_\lambda, s_\mu\rangle = \delta_{\lambda\mu}$.

Proof. Combine Corollary 7.10.6 and Lemma 7.9.2. $\square$

Linear-algebraic note. We say that the Schur functions form an *integral* orthonormal basis of Λ, since by Proposition 7.10.5 they actually generate $\Lambda_{\mathbb{Z}}$ as

an abelian group. In general it is a subtle question whether a vector space with a distinguished basis (in our case the monomial symmetric functions) and a positive definite symmetric scalar product has an integral orthonormal basis. For our situation such a basis is equivalent to the existence of an *integral* matrix A such that $A^t A = N$, where N is the transition matrix from m_λ to h_λ. We then say N is *integrally equivalent to the identity*. The next result (which is nothing more than standard linear algebra) identifies A as the Kostka matrix K. Note that in general if an integral orthonormal basis exists, then it is unique up to sign and order. This is because the transition matrix between two such bases must be both integral and orthogonal. It is easy to see that the only integral orthogonal matrices are signed permutation matrices.

7.12.3 Corollary. *Fix partitions $\mu, \nu \vdash n$. Then*

$$\sum_{\lambda \vdash n} K_{\lambda\mu} K_{\lambda\nu} = N_{\mu\nu} = \langle h_\mu, h_\nu \rangle,$$

where $K_{\lambda\mu}$ and $K_{\lambda\nu}$ denote Kostka numbers, and $N_{\mu\nu}$ is the number of $\mathbb{N}$-matrices A with $\mathrm{row}(A) = \mu$ *and* $\mathrm{col}(A) = \nu$.

Proof. Take the coefficient of $x^\mu y^\nu$ on both sides of (7.44). □

7.12.4 Corollary. *We have*

$$h_\mu = \sum_\lambda K_{\lambda\mu} s_\lambda. \tag{7.46}$$

In other words, if $M(u, v)$ denotes the transition matrix from the basis $\{v_\lambda\}$ to the basis $\{u_\lambda\}$ of Λ (so that $u_\lambda = \sum_\mu M(u, v)_{\lambda\mu} v_\mu$), then

$$M(h, s) = M(s, m)^t.$$

We give three proofs of this corollary, all essentially equivalent.

First Proof. Let $h_\mu = \sum_\lambda a_{\lambda\mu} s_\lambda$. By Corollary 7.12.2, we have $a_{\lambda\mu} = \langle h_\mu, s_\lambda \rangle$. Since $\langle h_\mu, m_\nu \rangle = \delta_{\mu\nu}$ by the definition (7.30) of the scalar product $\langle\ ,\ \rangle$, we have from (7.35) that $\langle h_\mu, s_\lambda \rangle = K_{\lambda\mu}$. □

Second Proof. Fix μ. Then

$$\begin{aligned} h_\mu &= \sum_A x^{\mathrm{col}(A)} \\ &= \sum_{(P,Q)} x^Q \qquad \text{by the RSK algorithm} \\ &= \sum_\lambda K_{\lambda\mu} \sum_Q x^Q \\ &= \sum_\lambda K_{\lambda\mu} s_\lambda, \end{aligned}$$

where (i) A ranges over all $\mathbb{N}$-matrices with $\text{row}(A) = \mu$, (ii) (P, Q) ranges over all pairs of SSYT of the same shape with $\text{type}(P) = \mu$, and (iii) Q ranges over all SSYT of shape λ. □

Third Proof. Take the coefficient of $m_\mu(x)$ on both sides of the identity

$$\sum_\lambda m_\lambda(x)h_\lambda(y) = \sum_\lambda s_\lambda(x)s_\lambda(y).$$

(The two sides are equal by (7.10) and (7.44).) □

The next corollary may be regarded as giving a generating function (with respect to the Schur functions s_λ) for the number f^λ of SYT of shape λ.

7.12.5 Corollary. We have

$$h_1^n = \sum_{\lambda \vdash n} f^\lambda s_\lambda. \tag{7.47}$$

Proof. Take the coefficient of $x_1x_2\cdots x_n$ on both sides of (7.44). To obtain a bijective proof, consider the RSK algorithm $A \xrightarrow{\text{RSK}} (P, Q)$ when $\text{col}(A) = \langle 1^n \rangle$. □

Finally we come to an identity already given in (7.43) but worth repeating here.

7.12.6 Corollary. *We have*

$$\sum_{\lambda \vdash n} (f^\lambda)^2 = n!.$$

Proof. Regard (7.47) as being in the variables $x = (x_1, x_2, \ldots)$, and take the coefficient of $x_1x_2\cdots x_n$ on both sides. To obtain a bijective proof (as mentioned before equation (7.43)) consider the RSK algorithm applied to $n \times n$ permutation matrices. □

7.13 Symmetry of the RSK Algorithm

The RSK algorithm has a number of remarkable symmetry properties. We will discuss only the most important such property in this section.

7.13.1 Theorem. *Let A be an $\mathbb{N}$-matrix of finite support, and suppose that $A \xrightarrow{\text{RSK}} (P, Q)$. Then $A^t \xrightarrow{\text{RSK}} (Q, P)$, where ^{t} denotes transpose.*

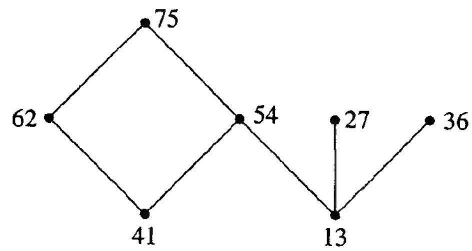
To prepare for the proof of this theorem, let $w_A = \binom{u}{v}$ be the two-line array associated to A. Hence $w_{A^t} = \binom{v}{u}_{\text{sorted}}$, i.e., sort the columns of $\binom{v}{u}$ so that the columns are weakly increasing in lexicographic order. It follows from Lemma 7.11.6 that we may assume u and v have no repeated elements [why?].

Given

$$w_A = \begin{pmatrix} u_1 & \cdots & u_n \\ v_1 & \cdots & v_n \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix},$$

where the u_i's and the v_j's are distinct, define the *inversion poset* $I = I(A) = I\binom{u}{v}$ as follows. The vertices of I are the columns of $\binom{u}{v}$. For notational convenience, we often denote a column ${a \atop b}$ as ab. Define $ab < cd$ in I if $a < c$ and $b < d$.

7.13.2 Example. Let $\binom{u}{v} = \begin{pmatrix} 1&2&3&4&5&6&7 \\ 3&7&6&1&4&2&5 \end{pmatrix}$. Then I is given by



Note that the number of incomparable pairs in I is just the number of inversions of the permutation v, whence the terminology "inversion poset."

The following lemma is an immediate consequence of the definition of $I(A)$.

7.13.3 Lemma. *The map $\varphi : I(A) \to I(A^t)$ defined by $\varphi(ab) = ba$ is an isomorphism of posets.*

Now given the inversion poset $I = I(A)$, define I_1 to be the set of minimal elements of I, then I_2 to be the set of minimal elements of $I - I_1$, then I_3 to be the set of minimal elements of $I - I_1 - I_2$, etc. For the poset of Example 7.13.2 we have $I_1 = \{13, 41\}$, $I_2 = \{27, 36, 54, 62\}$, $I_3 = \{75\}$. Note that since I_i is an antichain of I, its elements can be labeled

$$(u_{i1}, v_{i1}),\ (u_{i2}, v_{i2}),\ \ldots,\ (u_{in_i},\ v_{in_i}), \tag{7.48}$$

where $n_i = \#I_i$, such that

$$\begin{aligned} u_{i1} &< u_{i2} < \cdots < u_{in_i} \\ v_{i1} &> v_{i2} > \cdots > v_{in_i}. \end{aligned} \tag{7.49}$$

7.13.4 Lemma. *Let $I_1, \ldots, I_d$ be the (nonempty) antichains defined above, labeled as in (7.49). Let $A \xrightarrow{\text{RSK}} (P, Q)$. Then the first row of P is $v_{1n_1} v_{2n_2} \cdots v_{dn_d}$, while the first row of Q is $u_{11}u_{21} \cdots u_{d1}$. Moreover, if $(u_k, v_k) \in I_i$, then v_k is inserted into the i-th column of the first row of the tableau $P(k-1)$ in the RSK algorithm.*

Proof. Induction on n, the case $n = 1$ being trivial. Assume the assertion for $n - 1$, and let

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \\ v_1 & v_2 & \cdots & v_n \end{pmatrix}, \qquad \begin{pmatrix} \tilde{u} \\ \tilde{v} \end{pmatrix} = \begin{pmatrix} u_1 & u_2 & \cdots & u_{n-1} \\ v_1 & v_2 & \cdots & v_{n-1} \end{pmatrix}.$$

Let $(P(n-1), Q(n-1))$ be the tableaux obtained after inserting $v_1, \ldots, v_{n-1}$, and let the antichains $I_i' := I_i\binom{\tilde{u}}{\tilde{v}}$, $1 \le i \le e$ (where $e = d-1$ or $e = d$), be given by $(\tilde{u}_{i1}, \tilde{v}_{i1}), \ldots, (\tilde{u}_{im_i}, \tilde{v}_{im_i})$, where $\tilde{u}_{i1} < \cdots < \tilde{u}_{im_i}$ and $\tilde{v}_{i1} > \cdots > \tilde{v}_{im_i}$. By the induction hypothesis, the first row of $P(n-1)$ is $\tilde{v}_{1m_1}\tilde{v}_{2m_2}\cdots\tilde{v}_{em_e}$, while the first row of Q is $\tilde{u}_{11}\tilde{u}_{21}\cdots\tilde{u}_{e1}$. Now we insert v_n into $P(n-1)$. If $\tilde{v}_{im_i} > v_n$, then $I_i' \cup (u_n, v_n)$ is an antichain of $I\binom{u}{v}$. Hence $(u_n, v_n) \in I_i\binom{u}{v}$ if i is the *least* index for which $\tilde{v}_{im_i} > v_n$. If there is no such i, then (u_n, v_n) is the unique element of the antichain $I_d\binom{u}{v}$ of $I\binom{u}{v}$. These conditions mean that v_n is inserted into the i-th column of $P(n-1)$, as claimed. We start a new i-th column exactly when $v_n = v_{d1}$, in which case $u_n = u_{d1}$, so u_n is inserted into the i-th column of the first row of $Q(n-1)$, as desired. □

Proof of Theorem 7.13.1. If the antichain $I_i\binom{u}{v}$ is given by (7.48) such that (7.49) is satisfied, then by Lemma 7.13.3 the antichain $I_i\binom{v}{u}$ is just

$$(v_{im_i}, u_{im_i}), \ldots, (v_{i2}, u_{i2}), (v_{i1}, u_{i1}),$$

where

$$v_{im_i} < \cdots < v_{i2} < v_{i1}$$
$$u_{im_i} > \cdots > u_{i2} > u_{i1}.$$

Hence by Lemma 7.13.4, if $A^t \xrightarrow{\text{RSK}} (P', Q')$, then the first row of P' is $u_{11}u_{21}\cdots u_{d1}$, and the first row of Q' is $v_{im_1}v_{2m_2}\cdots v_{dm_d}$. Thus by Lemma 7.13.4, the first rows of P' and Q' agree with the first rows of Q and P, respectively.

When the RSK algorithm is applied to $\binom{u}{v}$, the element v_{ij}, $1 \le j < m_i$, gets bumped into the second row of P *before* the element v_{rs}, $1 \le s < m_r$, if and only if $u_{i,j+1} < u_{r,s+1}$. Let $\bar{P}$ and $\bar{Q}$ denote P and Q with their first rows removed. It follows that

$$\begin{aligned}\binom{a}{b} &:= \begin{pmatrix} u_{12}\cdots u_{1m_1} & u_{22}\cdots u_{2m_2} & \cdots u_{d2}\cdots u_{dm_d} \\ v_{11}\cdots v_{1,m_1-1} & v_{21}\cdots v_{2,m_2-1} & \cdots v_{d1}\cdots v_{d,m_d-1}\end{pmatrix}_{\text{sorted}} \\ &\xrightarrow{\text{RSK}} (\bar{P}, \bar{Q}).\end{aligned}$$

Similarly let $(\bar{P}', \bar{Q}')$ denote P' and Q' with their first rows removed. Applying the same argument to $\binom{v}{u}$ rather than $\binom{u}{v}$ yields

$$\begin{aligned}\binom{a'}{b'} &:= \begin{pmatrix} v_{1,m_1-1}\cdots v_{11} & v_{2,m_2-1}\cdots v_{21} & \cdots v_{d,m_d-1}\cdots v_{d1} \\ u_{1m_1}\quad\cdots u_{12} & u_{2m_2}\quad\cdots u_{22} & \cdots u_{dm_d}\quad\cdots u_{d2}\end{pmatrix}_{\text{sorted}} \\ &\xrightarrow{\text{RSK}} (\bar{P}', \bar{Q}').\end{aligned}$$

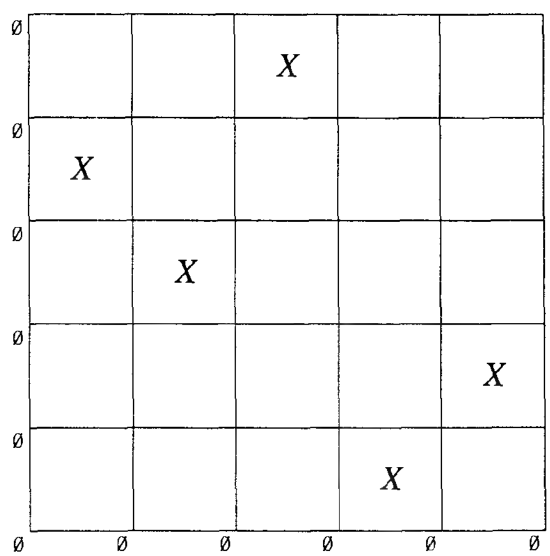
But $\binom{a}{b} = \binom{b'}{a'}_{\text{sorted}}$, so by induction on n (or on the number of rows) we have $(\bar{P}', \bar{Q}') = (\bar{Q}, \bar{P})$, and the proof follows. □

Second Proof (sketch). The above proof was somewhat mysterious and did not really "display" the symmetric nature of the RSK algorithm. We will describe an alternative "geometric" description of the RSK algorithm from which the symmetry property is obvious.

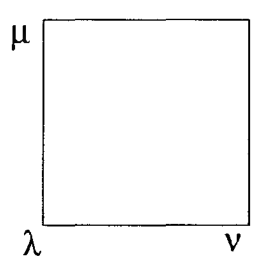
Given $w = w_1 \cdots w_n \in \mathfrak{S}_n$, construct an $n \times n$ square array with an X in the w_i-th square from the bottom of column i. For instance, if $w = 43512$ then we obtain

		X		
X				
	X			
				X
			X	
4	3	5	1	2

This is essentially the usual way of representing a permutation by a permutation matrix, except that we place the $(1, 1)$ entry at the bottom left instead of at the top left. We want to label each of the $(n+1)^2$ points that are corners of squares of our $n \times n$ array with a partition. We will write this partition just below and to the left of its corresponding point. Begin by labeling all points on the bottom row and left column with the empty partition $\emptyset$:



Suppose now that we have labeled all the corners of a square s except the upper right, say as follows:



Then label the upper right corner by the partition ρ defined by the following "local" rules (L1)–(L4):

(L1) Suppose that the square s does not contain an X, and that $\lambda = \mu = \nu$. Then define $\rho = \lambda$.
(L2) Suppose that s does not contain an X, and that $\lambda \subset \mu = \nu$. This automatically implies that $|\mu/\lambda| = 1$, so μ is obtained from λ by adding 1 to some part λ_i. Let ρ be obtained from μ by adding 1 to μ_{i+1}.
(L3) Suppose that s does not contain an X and that $\mu \neq \nu$. Define $\rho = \mu \cup \nu$, i.e., $\rho_i = \max(\mu_i, \nu_i)$.
(L4) Suppose that s contains an X. This automatically implies that $\lambda = \mu = \nu$. Let ρ be obtained from λ by adding 1 to λ_1.

Using these rules, we can uniquely label every square corner, one step at a time. The resulting array is called the *growth diagram* $\mathcal{G}_w$ of w. For our example $w = 43512$, we get the growth diagram

$$\begin{array}{ccccccccccc}
\emptyset & & 1 & & 11 & & 21 & & 211 & & 221 \\
 & & & & & X & & & & & \\
\emptyset & & 1 & & 11 & & 11 & & 111 & & 211 \\
 & X & & & & & & & & & \\
\emptyset & & \emptyset & & 1 & & 1 & & 11 & & 21 \\
 & & & X & & & & & & & \\
\emptyset & & \emptyset & & \emptyset & & \emptyset & & 1 & & 2 \\
 & & & & & & & & & X & \\
\emptyset & & \emptyset & & \emptyset & & \emptyset & & 1 & & 1 \\
 & & & & & & & X & & & \\
\emptyset & & \emptyset & & \emptyset & & \emptyset & & \emptyset & & \emptyset
\end{array} \tag{7.50}$$

It's easy to see that if a point p is labeled by λ, then the sum $|\lambda|$ of the parts of λ is equal to the number of X's in the quarter plane to the left and below p. In particular, if λ^i denotes the partition in row i (with the bottom row being row 0) and column n (the rightmost column), then $|\lambda^i| = i$. Moreover, it is immediate from the labeling procedure that $\emptyset = \lambda^0 \subset \lambda^1 \subset \cdots \subset \lambda^n$. Similarly, if μ^i denotes the partition in column i (with the leftmost column being 0) and row n (the top row), then $|\mu^i| = i$ and $\emptyset = \mu^0 \subset \mu^1 \subset \cdots \subset \mu^n$.

The chains $\emptyset = \lambda^0 \subset \lambda^1 \subset \cdots \subset \lambda^n$ and $\emptyset = \mu^0 \subset \mu^1 \subset \cdots \subset \mu^n$ correspond to standard tableaux P_w and Q_w, respectively (as explained in Proposition 7.10.3(a)). The main result concerning the geometric construction we have just described is the following.

7.13.5 Theorem. *The standard tableaux P_w and Q_w just described satisfy*

$$w \xrightarrow{\text{RSK}} (P_w, Q_w).$$

Proof (sketch). Let the partition appearing in row i and column j be $\nu(i, j)$. Thus for fixed j, we have

$$\emptyset = \nu(0, j) \subseteq \nu(1, j) \subseteq \cdots \subseteq \nu(n, j),$$

where $|\nu(i, j)/\nu(i-1, j)| = 0$ or 1. Let $T(i, j)$ be the tableau of shape $\nu(i, j)$ obtained by inserting k into the square $\nu(k, j)/\nu(k-1, j)$ when $0 \leq k < i$ and $|\nu(k, j)/\nu(k-1, j)| = 1$. For the array (7.50) the tableaux $T(i, j)$ are given by

∅	4	3 4	35 4 X	15 3 4	12 35 4
∅	4 X	3 4	3 4	1 3 4	12 3 4
∅	∅	3 X	3	1 3	12 3
∅	∅	∅	∅	1	12 X
∅	∅	∅	∅	1 X	1
∅	∅	∅	∅	∅	∅

We claim that the tableau $T(i, j)$ has the following alternative description: Let $(i_1, j_1), \ldots, (i_k, j_k)$ be the position of the x's to the left and below $T(i, j)$ (i.e., $i_r \leq i$ and $j_r \leq j$), labeled so that $j_1 < \cdots < j_k$. Then $T(i, j)$ is obtained by row inserting successively $i_1, i_2, \ldots, i_k$, beginning with an empty tableau. In symbols,

$$T(i, j) = ((\emptyset \leftarrow i_1) \leftarrow i_2) \leftarrow \cdots \leftarrow i_k.$$

The proof of the claim is by induction on $i + j$. The assertion is clearly true if $i = 0$ or $j = 0$, so that $T(i, j) = \emptyset$. If $i > 0$ and $j > 0$, then by the induction hypothesis we know that $T(i-1, j)$, $T(i, j-1)$, and $T(i-1, j-1)$ satisfy the desired conditions. One checks that $T(i, j)$ also satisfies these conditions by using the definition of $T(i, j)$ in terms of the local rules (L1)–(L4). There are thus four cases to check; we omit the rather straightforward details.

If we now take $i = n$, we see that

$$T(n, j) = ((\emptyset \leftarrow w_1) \leftarrow w_2) \leftarrow \cdots \leftarrow w_j, \tag{7.51}$$

where $w = w_1 w_2 \cdots w_n$. Thus $T(n, n)$ (which is the same as P_w) is indeed the

insertion tableau of w, while Q_w (which is defined by the sequence $\nu(n,0) \subset \nu(n,1) \subset \cdots \subset \nu(n,n)$) is by (7.51) just the recording tableau. This completes the proof of Theorem 7.13.5. $\square$

It is now almost trivial to give our second proof of Theorem 7.13.1. If we transpose the growth diagram G_w (i.e., reflect about the diagonal from the lower left to the upper right corner) then the symmetry of the local rules (L1)–(L4) with respect to transposition shows that we get simply the growth diagram $G_{w^{-1}}$. Hence $P_w = Q_{w^{-1}}$ and $Q_w = P_{w^{-1}}$, and the proof follows from Theorem 7.13.5.

Growth diagrams and their variants are powerful tools for understanding the RSK algorithm and related algorithms. For further infomation, see the Notes to this chapter, as well as Exercise 7.28(a).

Let us now consider some corollaries of the symmetry property given by Theorem 7.13.1.

7.13.6 Corollary. *Let A be an $\mathbb{N}$-matrix of finite support, and let $A \xrightarrow{\text{RSK}} (P, Q)$. Then A is symmetric (i.e., $A = A^t$) if and only if $P = Q$.*

Proof. Immediate from the fact that $A^t \xrightarrow{\text{RSK}} (Q, P)$. $\square$

7.13.7 Corollary. *Let $A = A^t$ and $A \xrightarrow{\text{RSK}} (P, P)$, and let $\alpha = (\alpha_1, \alpha_2, \ldots)$, where $\alpha_i \in \mathbb{N}$ and $\sum \alpha_i < \infty$. Then the map $A \mapsto P$ establishes a bijection between symmetric $\mathbb{N}$-matrices with* $\text{row}(A) = \alpha$ *and SSYTs of type α.*

Proof. Follows from Corollary 7.13.6 and Theorem 7.11.5. $\square$

7.13.8 Corollary. *We have*

$$\frac{1}{\prod_i (1 - x_i) \cdot \prod_{i<j} (1 - x_i x_j)} = \sum_\lambda s_\lambda(x), \tag{7.52}$$

summed over all $\lambda \in$ Par.

Proof. The coefficient of x^α on the left-hand side is the number of symmetric $\mathbb{N}$-matrices A with $\text{row}(A) = \alpha$ [why?], while the coefficient of x^α on the right-hand side is the number of SSYTs of type α. Now apply Corollary 7.13.7. $\square$

7.13.9 Corollary. *We have*

$$\sum_{\lambda \vdash n} f^\lambda = \#\{w \in \mathfrak{S}_n : w^2 = 1\},$$

the number of involutions in $\mathfrak{S}_n$ (discussed in Example 5.2.10).

Proof. Let $w \in \mathfrak{S}_n$ and $w \xrightarrow{\text{RSK}} (P, Q)$, where P and Q are SYT of the same shape $\lambda \vdash n$. The permutation matrix corresponding to w is symmetric if and only if $w^2 = 1$. By Theorem 7.13.1 this is the case if and only if $P = Q$, and the proof follows.

Alternatively, take the coefficient of $x_1 \cdots x_n$ on both sides of (7.52). □

Corollary 7.13.9 asserts that the total number of SYT of size n is equal to the number of involutions in $\mathfrak{S}_n$. The RSK algorithm provides a bijective proof.

Note that if $t(n)$ denotes the coefficient of $x_1 \cdots x_n$ on the left-hand side of (7.52), then Example 7.8.5 shows directly that

$$\sum_{n \geq 0} t(n) \frac{x^n}{n!} = \exp\left(x + \frac{x^2}{2}\right),$$

in agreement with (5.32).

7.14 The Dual RSK Algorithm

There is a variation of the RSK algorithm that is related to the product $\prod(1 + x_i y_j)$ in the same way that the RSK algorithm itself is related to $\prod(1 - x_i y_j)^{-1}$. We call this variation the *dual RSK algorithm* and denote it by $A \xrightarrow{\text{RSK}^*} (P, Q)$. The matrix A will now be a $(0, 1)$ matrix of finite support. Form the two-line array w_A just as before. The RSK* algorithm proceeds exactly like the RSK algorithm, except that an element i bumps the leftmost element $\geq i$, rather than the leftmost element $> i$. (In particular, RSK and RSK* agree for permutation matrices.) It follows that each row of P is *strictly* increasing.

7.14.1 Example. Let

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Then

$$w_A = \begin{pmatrix} 1 & 1 & 2 & 3 & 3 & 4 & 5 \\ 1 & 3 & 2 & 1 & 3 & 3 & 2 \end{pmatrix}.$$

The arrays $(P(1), Q(1)), \ldots, (P(7), Q(7))$, with $(P, Q) = (P(7), Q(7))$, obtained from RSK* are as follows:

$P(i)$	$Q(i)$
1	1
1 3	1 1
1 2 3	1 1 2
1 2 1 3	1 1 2 3
1 2 3 1 3	1 1 3 2 3
1 2 3 1 3 3	1 1 3 2 4 3
1 2 3 1 2 3 3	1 1 3 2 4 3 5

7.14.2 Theorem. *The RSK* algorithm is a bijection between* $(0, 1)$*-matrices* A *of finite support and pairs* (P, Q) *such that* P^t *(the transpose of* P*) and* Q *are SSYTs, with* $\mathrm{sh}(P) = \mathrm{sh}(Q)$*. Moreover,* $\mathrm{col}(A) = \mathrm{type}(P)$ *and* $\mathrm{row}(A) = \mathrm{type}(Q)$.

The proof of Theorem 7.14.2 is analogous to that of Theorem 7.11.5 and will be omitted.

Exactly as we obtained the Cauchy identity (7.44) from the ordinary RSK algorithm, we have the following result, known as the *dual Cauchy identity*.

7.14.3 Theorem. *We have*

$$\prod_{i,j}(1 + x_i y_j) = \sum_{\lambda} s_\lambda(x) s_{\lambda'}(y).$$

An important consequence of Theorem 7.14.3 is the evaluation of ωs_λ. First we need to see the effect of ω, acting on the y variables, on the product $\prod(1 + x_i y_j)$.

7.14.4 Lemma. *Let* ω_y *denote* ω *acting on the* y *variables only (so we regard the* x_i*'s as constants commuting with* ω*). Then*

$$\omega_y \prod(1 - x_i y_j)^{-1} = \prod(1 + x_i y_j).$$

Proof. We have

$$\begin{aligned}\omega_y \prod(1-x_iy_j)^{-1} &= \omega_y \sum_\lambda m_\lambda(x)h_\lambda(y) \qquad \text{(by Proposition 7.5.3)}\\ &= \sum_\lambda m_\lambda(x)e_\lambda(y) \qquad \text{(by Theorem 7.6.1)}\\ &= \prod(1+x_iy_j) \qquad \text{(by Proposition 7.4.3).}\end{aligned}$$

□

An alternative proof can be given by expanding the products $\prod(1-x_iy_j)^{-1}$ and $\prod(1+x_iy_j)$ in terms of the power sum symmetric functions (equations (7.20) and (7.21)) and applying Proposition 7.7.5.

7.14.5 Theorem. *For every* $\lambda \in \mathrm{Par}$ *we have*

$$\omega s_\lambda = s_{\lambda'}.$$

Proof. We have

$$\begin{aligned}\sum_\lambda s_\lambda(x)s_{\lambda'}(y) &= \prod(1+x_iy_j) \qquad \text{(by Theorem 7.14.3)}\\ &= \omega_y\prod(1-x_iy_j)^{-1} \qquad \text{(by Lemma 7.14.4)}\\ &= \omega_y\sum_\lambda s_\lambda(x)s_\lambda(y) \qquad \text{(by Theorem 7.12.1)}\\ &= \sum_\lambda s_\lambda(x)\omega_y\left(s_\lambda(y)\right).\end{aligned}$$

Take the coefficient of $s_\lambda(x)$ on both sides. Since the $s_\lambda(x)$'s are linearly independent, we obtain $s_{\lambda'}(y) = \omega_y\left(s_\lambda(y)\right)$, or just $s_{\lambda'} = \omega s_\lambda$. □

Later (Theorem 7.15.6) we will extend Theorem 7.14.5 to skew Schur functions.

After Proposition 7.7.5 we mentioned that the characteristic polynomial of $\omega : \Lambda^n \to \Lambda^n$ is equal to $(x^2-1)^{o(n)}(x-1)^{k(n)}$, where $o(n)$ is the number of odd conjugacy classes in $\mathfrak{S}_n$ and $k(n)$ is the number of self-conjugate partitions of n. In particular, the multiplicity of 1 as an eigenvalue exceeds the multiplicity of -1 by $k(n)$. This fact is also an immediate consequence of Theorem 7.14.5. For if $\lambda \neq \lambda'$ then ω transposes s_λ and $s_{\lambda'}$, accounting for one eigenvalue equal to 1 and one equal to -1. Left over are the $k(n)$ eigenvectors s_λ for which $\lambda = \lambda'$, with eigenvalue 1.

7.15 The Classical Definition of Schur Functions

Let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ and $w \in \mathfrak{S}_n$. As usual write $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$, and define

$$w(x^\alpha) = x_1^{\alpha_{w(1)}} \cdots x_n^{\alpha_{w(n)}}.$$

Now define

$$a_\alpha = a_\alpha(x_1, \ldots, x_n) = \sum_{w \in \mathfrak{S}_n} \varepsilon_w w(x^\alpha), \tag{7.53}$$

where

$$\varepsilon_w = \begin{cases} 1 & \text{if } w \text{ is an even permutation} \\ -1 & \text{if odd.} \end{cases}$$

(Thus $\varepsilon_w = \varepsilon_{\rho(w)}$, as defined in (7.19).) Note that the right-hand side of equation (7.53) is just the expansion of a determinant, namely,

$$a_\alpha = \det\left(x_i^{\alpha_j}\right)_{i,j=1}^n .$$

Note also that a_α is skew-symmetric, i.e., $w(a_\alpha) = \varepsilon_w a_\alpha$, so $a_\alpha = 0$ unless all the α_i's are distinct. Hence assume that $\alpha_1 > \alpha_2 > \cdots > \alpha_n \geq 0$, so $\alpha = \lambda + \delta$, where $\lambda \in \text{Par}$, $\ell(\lambda) \leq n$, and $\delta = \delta_n = (n-1, n-2, \ldots, 0)$. Since $\alpha_j = \lambda_j + n - j$, we get

$$a_\alpha = a_{\lambda+\delta} = \det\left(x_i^{\lambda_j+n-j}\right)_{i,j=1}^n . \tag{7.54}$$

For instance,

$$a_{421} = a_{211+210} = \begin{vmatrix} x_1^4 & x_1^2 & x_1^1 \\ x_2^4 & x_2^2 & x_2^1 \\ x_3^4 & x_3^2 & x_3^1 \end{vmatrix}.$$

Note in particular that

$$a_\delta = \det\left(x_i^{n-j}\right) = \prod_{1 \leq i < j \leq n} (x_i - x_j), \tag{7.55}$$

the *Vandermonde determinant*.

If for some $i \neq j$ we put $x_i = x_j$ in a_α, then because a_α is skew-symmetric (or because the i-th row and j-th row of the determinant (7.54) become equal), we obtain 0. Hence a_α is divisible by $x_i - x_j$ and thus by a_δ (in the ring $\mathbb{Z}[x_1, \ldots, x_n]$).

Thus $a_\alpha/a_\delta \in \mathbb{Z}[x_1, \ldots, x_n]$. Moreover, since a_α and a_δ are skew-symmetric, the quotient is symmetric, and is clearly homogeneous of degree $|\alpha| - |\delta| = |\lambda|$. In other words, $a_\alpha/a_\delta \in \Lambda_n^{|\lambda|}$. (The quotient a_α/a_δ is called a *bialternant*.) It is therefore natural to ask for the relation between a_α/a_δ and the symmetric functions we have already considered. The answer is a fundamental result in the theory of symmetric functions.

7.15.1 Theorem. *We have*

$$a_{\lambda+\delta}/a_\delta = s_\lambda(x_1, \ldots, x_n).$$

Proof. There are many proofs of this result. We give one that can be extended to give an important result on skew Schur functions (Theorem 7.15.4).

Applying ω to (7.46) and replacing λ by λ' yields

$$e_\mu = \sum_\lambda K_{\lambda'\mu} s_\lambda.$$

Since the matrix $(K_{\lambda'\mu})$ is invertible, it suffices to show that

$$e_\mu(x_1, \ldots, x_n) = \sum_\lambda K_{\lambda'\mu} \frac{a_{\lambda+\delta}}{a_\delta},$$

or equivalently (always working with n variables),

$$a_\delta e_\mu = \sum_\lambda K_{\lambda'\mu} a_{\lambda+\delta}. \tag{7.56}$$

Since both sides of (7.56) are skew-symmetric, it is enough to show that the coefficient of $x^{\lambda+\delta}$ in $a_\delta e_\mu$ is $K_{\lambda'\mu}$. We multiply a_δ by e_μ by successively multiplying by $e_{\mu_1}, e_{\mu_2}, \ldots$. Each partial product $a_\delta e_{\mu_1} \cdots e_{\mu_k}$ is skew-symmetric, so any term $x_1^{i_1} \cdots x_n^{i_n}$ appearing in $a_\delta e_{\mu_1} \cdots e_{\mu_k}$ has all exponents i_j *distinct*. When we multiply such a term $x_1^{i_1} \cdots x_n^{i_n}$ by a term $x_{m_1} \cdots x_{m_j}$ from $e_{\mu_{k+1}}$ (so $j = \mu_{k+1}$), either two exponents become equal or the exponents maintain their relative order. If two exponents become equal, then that term disappears from $a_\delta e_{\mu_1} \cdots e_{\mu_{k+1}}$. Hence to get the term $x^{\lambda+\delta}$, we must start with the term x^δ in a_δ and successively multiply by a term x^{α^1} of e_{μ_1}, then x^{α^2} of e_{μ_2}, etc., keeping the exponents strictly decreasing. The number of ways to do this is the coefficient of $x^{\lambda+\delta}$ in $a_\delta e_\mu$.

Given the terms $x^{\alpha^1}, x^{\alpha^2}, \ldots$ as above, define an SSYT $T = T(\alpha^1, \alpha^2, \ldots)$ as follows: Column j of T contains an i if the variable x_j occurs in x^{α^i} (i.e., the j-th coordinate of α^i is equal to 1). For example, suppose $n = 4$, $\lambda = 5332$, $\lambda' = 44311$, $\lambda + \delta = 8542$, $\mu = 3222211$, $x^{\alpha^1} = x_1x_2x_3$, $x^{\alpha^2} = x_1x_2$, $x^{\alpha^3} = x_3x_4$,

$x^{\alpha^4} = x_1x_2$, $x^{\alpha^5} = x_1x_4$, $x^{\alpha^6} = x_1$, $x^{\alpha^7} = x_3$. Then T is given by

$$\begin{array}{llll} 1 & 1 & 1 & 3 \\ 2 & 2 & 3 & 5 \\ 4 & 4 & 7 & \\ 5 & & & \\ 6 & & & \end{array}$$

It is easy to see that the map $(\alpha^1, \alpha^2, \ldots) \mapsto T(\alpha^1, \alpha^2, \ldots)$ gives a bijection between ways of building up the term $x^{\lambda+\delta}$ from x^δ (according to the rules above) and SSYT of shape λ' and type μ, so the proof follows. □

From the combinatorial definition of Schur functions it is clear that $s_\lambda(x_1, \ldots, x_n) = 0$ if $\ell(\lambda) > n$. Since by Proposition 7.8.2(b) we have $\dim \Lambda_n = \#\{\lambda \in \mathrm{Par} : \ell(\lambda) \leq n\}$, it follows that the set $\{s_\lambda(x_1, \ldots, x_n) : \ell(\lambda) \leq n\}$ is a basis for Λ_n. (This also follows from a simple extension of the proof of Corollary 7.10.6.) We define on Λ_n a scalar product $\langle\ ,\ \rangle_n$ by requiring that $\{s_\lambda(x_1, \ldots, x_n)\}$ is an orthonormal basis. If $f, g \in \Lambda$, then we write $\langle f, g\rangle_n$ as short for $\langle f(x_1, \ldots, x_n), g(x_1, \ldots, x_n)\rangle_n$. Thus

$$\langle f, g\rangle = \langle f, g\rangle_n,$$

provided that every monomial appearing in f involves at most n distinct variables, e.g., if $\deg f \leq n$.

7.15.2 Corollary. *If $f \in \Lambda_n$, $\ell(\lambda) \leq n$, and $\delta = (n-1, n-2, \ldots, 1, 0)$, then*

$$\langle f, s_\lambda\rangle_n = [x^{\lambda+\delta}]a_\delta f,$$

the coefficient of $x^{\lambda+\delta}$ in $a_\delta f$.

Proof. All functions will be in the variables $x_1, \ldots, x_n$. Let $f = \sum_{\ell(\lambda)\leq n} c_\lambda s_\lambda$. Then by Theorem 7.15.1 we have

$$a_\delta f = \sum_{\ell(\lambda)\leq n} c_\lambda a_{\lambda+\delta},$$

so

$$\langle f, s_\lambda\rangle_n = c_\lambda = [x^{\lambda+\delta}]a_\delta f.$$ □

For instance, we have

$$\left\langle a_\delta^{2k}, s_\lambda\right\rangle_n = [x^{\lambda+\delta}]a_\delta^{2k+1}, \tag{7.57}$$

for $\ell(\lambda) \leq n$. It is an interesting problem (not completely solved) to compute the numbers (7.57); for further information on the case $k = 1$, see Exercise 7.37.

Let us now consider a "skew generalization" of Theorem 7.15.1. We continue to work in the n variables $x_1, \ldots, x_n$. For any $\lambda, \nu \in$ Par, $\ell(\lambda) \leq n$, $\ell(\nu) \leq n$, consider the expansion

$$s_\nu e_\mu = \sum_\lambda L^{\lambda'}_{\nu'\mu} s_\lambda,$$

or equivalently (multiplying by a_δ),

$$a_{\nu+\delta} e_\mu = \sum_\lambda L^{\lambda'}_{\nu'\mu} a_{\lambda+\delta}. \tag{7.58}$$

Arguing as in the proof of Theorem 7.15.1 shows that $L^{\lambda'}_{\nu'\mu}$ is equal to the number of ways to write

$$\lambda + \delta = \nu + \delta + \alpha^1 + \alpha^2 + \cdots + \alpha^k,$$

where $\ell(\mu) = k$, each α^i is a $(0, 1)$-vector with μ_i 1's, and each partial sum $\nu + \delta + \alpha^1 + \cdots + \alpha^i$ has strictly decreasing coordinates. Define a skew SSYT $T = T_{\lambda'/\nu'}(\alpha^1, \ldots, \alpha^k)$ of shape λ'/ν' and type μ by the condition that i appears in column j of T if the j-th coordinate of α^i is a 1. This establishes a bijection which shows that $L^{\lambda'}_{\nu'\mu}$ is equal to the skew Kostka number $K_{\lambda'/\nu',\mu}$, the number of skew SSYTs of shape λ'/ν' and type μ (see equation (7.36)). (If $\nu' \not\subseteq \lambda'$ then this number is 0.)

7.15.3 Corollary. *We have*

$$s_\nu e_\mu = \sum_\lambda K_{\lambda'/\nu',\mu} s_\lambda. \tag{7.59}$$

Proof. Divide (7.58) by a_δ, and let $n \to \infty$. $\square$

It is now easy to establish a fundamental property of skew Schur functions.

7.15.4 Theorem. *For any $f \in \Lambda$, we have*

$$\langle f s_\nu, s_\lambda \rangle = \langle f, s_{\lambda/\nu} \rangle.$$

In other words, the two linear transformations $M_\nu : \Lambda \to \Lambda$ and $D_\nu : \Lambda \to \Lambda$ defined by $M_\nu f = s_\nu f$ and $D_\nu s_\lambda = s_{\lambda/\nu}$ are adjoint with respect to the scalar product $\langle\ ,\ \rangle$. In particular,

$$\langle s_\mu s_\nu, s_\lambda \rangle = \langle s_\mu, s_{\lambda/\nu} \rangle. \tag{7.60}$$

Proof. Apply ω to (7.59) and replace ν by ν' and λ by λ'. We obtain

$$s_\nu h_\mu = \sum_\lambda K_{\lambda/\nu,\mu} s_\lambda.$$

Hence

$$\langle s_\nu h_\mu, s_\lambda \rangle = K_{\lambda/\nu,\mu} = \langle h_\mu, s_{\lambda/\nu} \rangle, \tag{7.61}$$

by (7.36) and the fact that $\langle h_\mu, m_\rho \rangle = \delta_{\mu\rho}$ by definition of $\langle\ ,\ \rangle$. But equation (7.61) is linear in h_μ, so since $\{h_\mu\}$ is a basis for Λ, the proof follows. □

7.15.5 Example. We have $s_1 s_{31} = s_{41} + s_{32} + s_{311}$ and $s_1 s_{22} = s_{32} + s_{221}$. No other product $s_1 s_\mu$ involves s_{32}. It follows that $s_{32/1} = s_{22} + s_{31}$. For a generalization, see Corollary 7.15.9.

We can now give the generalization of Theorem 7.14.5 to skew Schur functions.

7.15.6 Theorem. *For any* $\lambda, \nu \in \mathrm{Par}$ *we have* $\omega s_{\lambda/\nu} = s_{\lambda'/\nu'}$.

Proof. By Proposition 7.9.5 and equation (7.60) we have

$$\langle \omega(s_\mu s_\nu), \omega s_\lambda \rangle = \langle \omega s_\mu, \omega s_{\lambda/\nu} \rangle.$$

Hence by Theorem 7.14.5 we get

$$\langle s_{\mu'} s_{\nu'}, s_{\lambda'} \rangle = \langle s_{\mu'}, \omega s_{\lambda/\nu} \rangle. \tag{7.62}$$

On the other hand, substituting λ', μ', ν' for λ, μ, ν respectively in (7.60) yields

$$\langle s_{\mu'} s_{\nu'}, s_{\lambda'} \rangle = \langle s_{\mu'}, s_{\lambda'/\nu'} \rangle. \tag{7.63}$$

From (7.62) and (7.63) there follows $\omega s_{\lambda/\nu} = s_{\lambda'/\nu'}$. □

The integer $\langle s_\lambda, s_\mu s_\nu \rangle = \langle s_{\lambda/\nu}, s_\mu \rangle = \langle s_{\lambda/\mu}, s_\nu \rangle$ is denoted $c^\lambda_{\mu\nu}$ and is called a *Littlewood–Richardson coefficient*. Thus

$$\begin{aligned} s_\mu s_\nu &= \sum_\lambda c^\lambda_{\mu\nu} s_\lambda \\ s_{\lambda/\nu} &= \sum_\mu c^\lambda_{\mu\nu} s_\mu \\ s_{\lambda/\mu} &= \sum_\nu c^\lambda_{\mu\nu} s_\nu. \end{aligned} \tag{7.64}$$

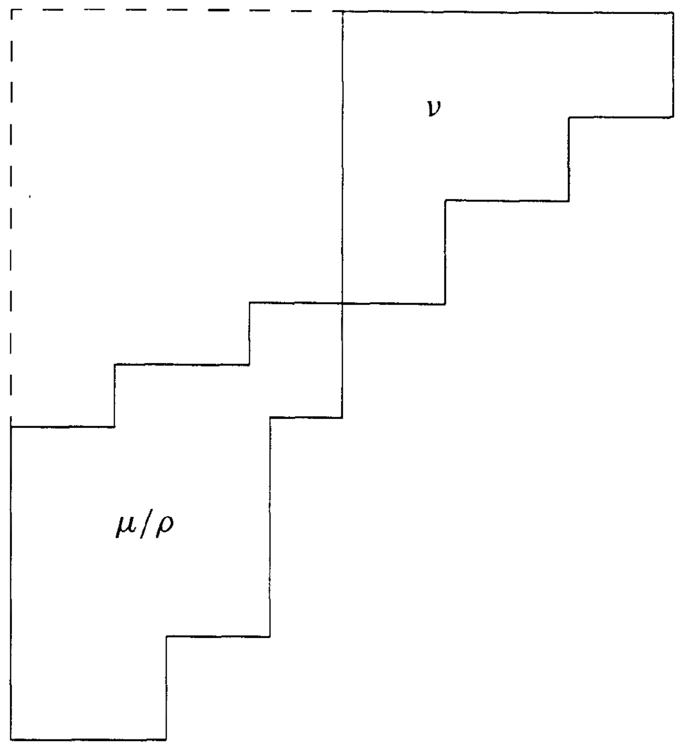


Figure 7-5. A skew shape.

Note that the seemingly more general $\langle s_{\lambda/\nu}, s_{\mu/\rho}\rangle$ is itself a Littlewood–Richardson coefficient, since $\langle s_{\lambda/\nu}, s_{\mu/\rho}\rangle = \langle s_\lambda, s_\nu s_{\mu/\rho}\rangle$ and $s_\nu s_{\mu/\rho}$ is just a skew Schur function, as Figure 7-5 (together with the combinatorial definition of Schur functions) makes evident. More generally, any product of skew Schur functions is a skew Schur function.

A central result in the theory of symmetric functions, called the *Littlewood–Richardson rule*, gives a combinatorial interpretation of the Littlewood–Richardson coefficient $c^\lambda_{\mu\nu}$. We will defer the statement and proof of the Littlewood–Richardson rule to Appendix 1 (Section A1.3). Here we consider the much easier special case when $\mu = (n)$, the partition with a single part equal to n. To state this result, known as *Pieri's rule*, define a *horizontal strip* to be a skew shape λ/ν with no two squares in the same column. Thus an SSYT of shape μ/ρ with largest part at most m may be regarded as a sequence $\rho = \mu^0 \subseteq \mu^1 \subseteq \cdots \subseteq \mu^m = \mu$ of partitions such that each skew shape μ^i/μ^{i-1} is a horizontal strip. (Simply insert i into each square of μ^i/μ^{i-1}.)

7.15.7 Theorem. *We have*

$$s_\nu s_n = \sum_\lambda s_\lambda, \tag{7.65}$$

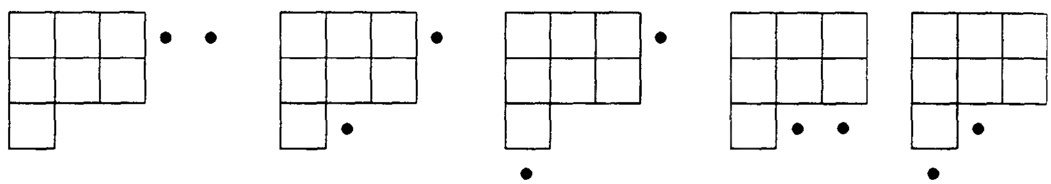
summed over all partitions λ such that λ/ν is a horizontal strip of size n.

Proof. We have $\langle s_\nu s_n, s_\lambda\rangle = \langle s_n, s_{\lambda/\nu}\rangle = \langle h_n, s_{\lambda/\nu}\rangle = K_{\lambda/\nu,n}$. Clearly by definition of $K_{\lambda/\nu,n}$ we have

$$K_{\lambda/\nu,n} = \begin{cases} 1 & \text{if } \lambda/\nu \text{ is a horizontal strip of size } n \\ 0 & \text{otherwise,} \end{cases}$$

and the proof follows. $\square$

7.15.8 Example. Let $\nu = 331$ and $n = 2$. The ways of adding a horizontal strip of size 2 to the shape 331 are given by



Hence

$$s_{331}s_2 = s_{531} + s_{432} + s_{4311} + s_{333} + s_{3321}.$$

Note that by applying ω to (7.65) we get a dual version of Pieri's rule. Namely, defining a *vertical strip* in the obvious way, we have

$$s_\nu s_{1^n} = s_\nu e_n = \sum_\lambda s_\lambda,$$

summed over all partitions λ for which λ/ν is a vertical strip of size n.

We also have as an immediate consequence of (7.60) and Pieri's rule (Theorem 7.15.7) the following skew version of Pieri's rule.

7.15.9 Corollary. *We have*

$$s_{\lambda/n} = \sum_\nu s_\nu,$$

where ν ranges over all partitions $\nu \subseteq \lambda$ for which λ/ν is a horizontal strip of size n.

The proof we have given of Pieri's rule is rather indirect, but Pieri's rule is actually a simple combinatorial statement that deserves a direct bijective proof. Let $\mathcal{T}^\alpha_{\nu,n}$ be the set of all pairs (T, T') of SSYTs such that $\mathrm{sh}(T) = \nu$, $\mathrm{sh}(T') = (n)$, and $\mathrm{type}(T) + \mathrm{type}(T') = \alpha$. Similarly, let $\mathcal{T}^\alpha_\lambda$ be the set of all SSYTs T such that $\mathrm{sh}(T) = \lambda$ and $\mathrm{type}(T) = \alpha$. Pieri's rule asserts that

$$\#\mathcal{T}^\alpha_{\nu,n} = \#\left(\bigcup_\lambda \mathcal{T}^\alpha_\lambda\right),$$

where λ ranges over all partitions such that λ/ν is a horizontal strip of size n. Thus

we seek a bijection

$$\varphi : \mathcal{T}^{\alpha}_{\nu,n} \to \bigcup_{\lambda} \mathcal{T}^{\alpha}_{\lambda}.$$

Let $T' = a_1 a_2 \cdots a_n$. It is not difficult to show (using Lemma 7.11.2) that φ is given just by iterated row insertion:

$$\varphi(T, T') = ((T \leftarrow a_1) \leftarrow a_2) \leftarrow \cdots \leftarrow a_n.$$

A further avatar of Theorem 7.15.4 is the following. Let $\Lambda(x)$ and $\Lambda(y)$ denote the rings of symmetric functions in the variables $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots)$, respectively. Denote by $\Lambda(x) \otimes \Lambda(y)$ the ring of formal power series (over $\mathbb{Q}$) in x and y of bounded degree that are symmetric in the x variables and symmetric in the y variables. In other words, if $f(x_1, x_2, \ldots; y_1, y_2, \ldots) \in \Lambda(x) \otimes \Lambda(y)$ and if u and v are both permutations of $\mathbb{P}$, then

$$f(x_1, x_2, \ldots; y_1, y_2, \ldots) = f(x_{u(1)}, x_{u(2)}, \ldots; y_{v(1)}, y_{v(2)}, \ldots).$$

It is clear that if $\{b_\mu(x)\}$ is a basis for $\Lambda(x)$ and $\{c_\nu(y)\}$ for $\Lambda(y)$, then $\{b_\mu(x)c_\nu(y)\}$ is a basis for $\Lambda(x) \otimes \Lambda(y)$. The ring $\Lambda(x, y)$ of formal power series of bounded degree that are symmetric in the x and y variables *together* is a subalgebra of $\Lambda(x) \otimes \Lambda(y)$. Of course the containment is proper; for instance, if $f(x) \in \Lambda(x)$ and $\deg f > 0$, then $f(x) \in \Lambda(x) \otimes \Lambda(y)$ but $f(x) \notin \Lambda(x, y)$. If $\{b_\lambda(x)\}$ is a basis for $\Lambda(x)$ then $\{b_\lambda(x, y)\}$ is a basis for $\Lambda(x, y)$, where $b_\lambda(x, y)$ denotes the symmetric function b_λ in the variables $x_1, x_2, \ldots$ and $y_1, y_2, \ldots$. It is now natural to ask how to expand $s_\lambda(x, y)$ in terms of the basis $\{s_\mu(x)s_\nu(y)\}$ of $\Lambda(x) \otimes \Lambda(y)$. Consider an ordered alphabet $A = \{1 < 2 < \cdots < 1' < 2' < \cdots\}$. If T is an SSYT of shape λ with respect to this alphabet, then define

$$(xy)^T = x_1^{\#(1)} x_2^{\#(2)} \cdots y_1^{\#(1')} y_2^{\#(2')} \cdots,$$

where $\#(a)$ denotes the number of occurrences of a in T. Thus from the combinatorial definition of s_λ (Definition 7.10.1), we have

$$s_\lambda(x, y) = \sum_T (xy)^T,$$

where T ranges over all SSYT of shape λ in the alphabet A. Now the part of T occupied by $1, 2, \ldots$ is just an SSYT of some shape $\mu \subseteq \lambda$, while the part of T occupied by $1', 2', \ldots$ is a skew SSYT of shape λ/μ. From this observation there follows

$$\begin{aligned} s_\lambda(x, y) &= \sum_{\mu \subseteq \lambda} s_\mu(x) s_{\lambda/\mu}(y) \\ &= \sum_{\mu \subseteq \lambda} s_\mu(x) \sum_\nu c^\lambda_{\mu\nu} s_\nu(y) \\ &= \sum_{\mu,\nu} c^\lambda_{\mu\nu} s_\mu(x) s_\nu(y), \end{aligned} \tag{7.66}$$

which gives us the desired expansion.

NOTE (for algebraists). Define $\Delta : \Lambda \to \Lambda(x) \otimes \Lambda(y)$ by $\Delta f = f(x, y)$. This operation makes the space Λ into a *coalgebra*, and together with the usual algebra structure on Λ it forms a *bialgebra*. If we take $1 \in \Lambda$ to be a unit and the map $f \mapsto f(0, 0, \ldots)$ to be a counit, then we get a *Hopf algebra*. Moreover, the scalar product on Λ is compatible with the bialgebra structure, in the sense that

$$\langle \Delta f, g(x)h(y) \rangle = \langle f, gh \rangle. \tag{7.67}$$

Here the first scalar product takes place in $\Lambda(x) \otimes \Lambda(y)$, where the elements $s_\mu(x)s_\nu(y)$ form an orthonormal basis. The second scalar product is just the usual one on Λ.

7.16 The Jacobi–Trudi Identity

In this section we will expand the Schur functions in terms of the complete symmetric functions. In effect we are computing the inverse to the Kostka matrix $(K_{\lambda\mu})$. Note that expanding Schur functions in terms of h_λ's is equivalent to expanding them in terms of e_λ's, for if $s_\lambda = \sum_\mu t_{\lambda\mu} h_\mu$, then applying ω yields $s_{\lambda'} = \sum_\mu t_{\lambda\mu} e_\mu$.

The main result of this section, known as the *Jacobi–Trudi identity*, expresses s_λ (in fact, $s_{\lambda/\mu}$) as a determinant whose entries are h_i's. Each term of the expansion of this determinant is thus of the form $\pm h_\nu$, so we get our desired expansion. The actual coefficient of h_ν must be obtained by collecting terms.

7.16.1 Theorem. *Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ and $\mu = (\mu_1, \ldots, \mu_n) \subseteq \lambda$. Then*

$$s_{\lambda/\mu} = \det\bigl(h_{\lambda_i - \mu_j - i + j}\bigr)_{i,j=1}^n , \tag{7.68}$$

where we set $h_0 = 1$ and $h_k = 0$ for $k < 0$.

First Proof. Our first proof will be a direct application of Theorem 2.7.1, in which we evaluated a determinant combinatorially by constructing an involution that canceled out all unwanted terms. Indeed, the Jacobi–Trudi identity is perhaps the archetypal application of Theorem 2.7.1.

In Theorem 2.7.1, take $\alpha_j = \lambda_j + n - j$, $\beta_i = \mu_i + n - i$, $\gamma_i = \infty$ (more precisely, take $\gamma_i = N$ and let $N \to \infty$), and $\delta_j = 1$. The function $h(\alpha_j - \beta_i; \gamma_i, \delta_j)$ appearing in Theorem 2.7.1 is just the complete symmetric function $h_{\lambda_j - \mu_i - j + i}$. Thus the determinant appearing in Theorem 2.7.1 becomes (after interchanging the roles of i and j) the right-hand side of equation (7.68).

Therefore by Theorem 2.7.1 it remains to show that $B(\alpha, \beta, \gamma, \delta) = s_{\lambda/\mu}$. In other words, given a nonintersecting n-path $\mathbf{L}$ in $\mathcal{B}(\alpha, \beta, \gamma, \delta)$, we need to associate (in a bijective fashion) a skew SSYT T of shape λ/μ such that the weight of $\mathbf{L}$ is equal to $x^{\mathrm{type}(T)}$. Actually, we associate a *reverse* SSYT T, which by Proposition 7.10.4 does not make any difference. If the horizontal steps of the path from $(\mu_i + n - i, \infty)$ to $(\lambda_i + n - i, 1)$ occur at heights $a_1 \geq a_2 \geq \cdots \geq a_{\lambda_i - \mu_i}$, then let $a_1, a_2, \ldots, a_{\lambda_i - \mu_i}$ be the ith row of T. A little thought shows that this establishes the desired bijection. $\square$

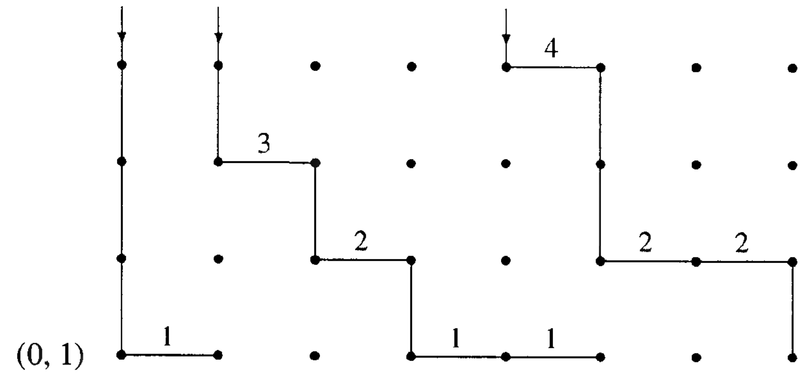


Figure 7-6. Nonintersecting lattice paths corresponding to an SSYT of shape 541/2.

As an example of the above bijection, take **L** as in Figure 7-6. Then

$$T = \begin{array}{ccccc} & & 4 & 2 & 2 \\ 3 & 2 & 1 & 1 & \\ 1 & & & & \end{array}.$$

Second Proof. Though our first proof was a very elegant combinatorial argument, it is also worthwhile to give a purely algebraic proof. Let $c^{\lambda}_{\mu\nu} = \langle s_\lambda, s_\mu s_\nu \rangle$, so

$$s_\mu s_\nu = \sum_\lambda c^{\lambda}_{\mu\nu} s_\lambda, \qquad s_{\lambda/\mu} = \sum_\nu c^{\lambda}_{\mu\nu} s_\nu.$$

Then

$$\begin{aligned} \sum_\lambda s_{\lambda/\mu}(x) s_\lambda(y) &= \sum_{\lambda,\nu} c^{\lambda}_{\mu\nu} s_\nu(x) s_\lambda(y) \\ &= \sum_\nu s_\nu(x) s_\mu(y) s_\nu(y) \\ &= s_\mu(y) \sum_\nu h_\nu(x) m_\nu(y). \end{aligned}$$

Let $y = (y_1, \ldots, y_n)$. Multiplying by $a_\delta(y)$ gives

$$\begin{aligned} \sum_\lambda s_{\lambda/\mu}(x) a_{\lambda+\delta}(y) &= \left(\sum_\nu h_\nu(x) m_\nu(y) \right) a_{\mu+\delta}(y) \\ &= \left(\sum_{\alpha \in \mathbb{N}^n} h_\alpha(x) y^\alpha \right) \left(\sum_{w \in \mathfrak{S}_n} \varepsilon_w y^{w(\mu+\delta)} \right) \\ &= \sum_{w \in \mathfrak{S}_n} \sum_\alpha \varepsilon_w h_\alpha(x) y^{\alpha + w(\mu+\delta)}. \end{aligned}$$

Now take the coefficient of $y^{\lambda+\delta}$ on both sides (so we are looking at terms where $\lambda + \delta = \alpha + w(\mu + \delta)$). We get

$$s_{\lambda/\mu}(x) = \sum_{w \in \mathfrak{S}_n} \varepsilon_w h_{\lambda+\delta-w(\mu+\delta)}(x) \tag{7.69}$$

$$= \det\left(h_{\lambda_i-\mu_j-i+j}(x)\right)_{i,j=1}^{n}.$$

If we substitute λ'/μ' for λ/μ in (7.68) and apply the automorphism ω, then we obtain the expansion of $s_{\lambda/\mu}$ in terms of the elementary symmetric functions, namely:

7.16.2 Corollary. *Let $\mu \subseteq \lambda$ with $\lambda_1 \le n$. Then*

$$s_{\lambda/\mu} = \det\left(e_{\lambda'_i-\mu'_j-i+j}\right)_{i,j=1}^{n}. \tag{7.70}$$

Equation (7.70) is known as the *dual Jacobi–Trudi identity*.

Recall (Proposition 7.8.4) that the exponential specialization ex satisfies

$$\mathrm{ex}(f) = \sum_{n\ge 0} [x_1x_2\cdots x_n] f \frac{t^n}{n!}.$$

Let $\mathrm{ex}_1(f) = \mathrm{ex}(f)_{t=1}$, provided this number is defined. In particular, if $|\lambda/\mu| = N$ then

$$\mathrm{ex}_1(s_{\lambda/\mu}) = \frac{f^{\lambda/\mu}}{N!},$$

where $f^{\lambda/\mu}$ is the number of SYT of shape λ/μ.

7.16.3 Corollary. *Let $|\lambda/\mu| = N$ and $\ell(\lambda) \le n$. Then*

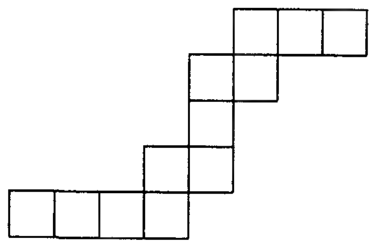
$$f^{\lambda/\mu} = N!\, \det\left(\frac{1}{(\lambda_i - \mu_j - i + j)!}\right)_{i,j=1}^{n}. \tag{7.71}$$

Proof. Apply ex_1 to the Jacobi–Trudi identity (equation (7.68)). Since $\mathrm{ex}_1(h_m) = 1/m!$ by (7.27), the proof follows. □

While it is certainly possible to prove Corollary 7.16.3 directly, our proof shows that it is just a specialization of the Jacobi–Trudi identity. When $\mu = \emptyset$ the determinant appearing in (7.71) can be explicitly evaluated (e.g., by induction and a clever use of row and column operations), thereby giving an explicit formula for f^{λ}. We will defer this formula to Corollary 7.21.6 and equation (7.113), where we give two less computational proofs.

7.17 The Murnaghan–Nakayama Rule

We have succeeded in expressing the Schur functions in terms of the bases m_λ, h_λ, and e_λ. In this section we consider the power sum symmetric functions p_λ. A skew shape λ/μ is *connected* if the interior of the diagram of λ/μ, regarded as a union of *solid* squares, is a connected (open) set. For instance, the shape $21/1$ is *not* connected. A *border strip* (or *rim hook* or *ribbon*) is a connected skew shape with no 2×2 square. An example of a border strip is $86554/5443$, whose diagram is



Given positive integers $a_1, \ldots, a_k$, there is a unique border strip λ/μ (up to translation) with a_i squares in row i (i.e., $a_i = \lambda_i - \mu_i$). It follows that the number of border strips of size n (up to translation) is 2^{n-1}, the number of compositions of n. Define the *height* $\mathrm{ht}(B)$ of a border strip B to be one less than its number of rows. The next result shows the connection between border strips and symmetric functions.

7.17.1 Theorem. *For any $\mu \in$ Par and $r \in \mathbb{N}$ we have*

$$s_\mu p_r = \sum_\lambda (-1)^{\mathrm{ht}(\lambda/\mu)} s_\lambda, \tag{7.72}$$

summed over all partitions $\lambda \supseteq \mu$ for which λ/μ is a border strip of size r.

Proof. Let $\delta = (n-1, n-2, \ldots, 0)$, and let all functions be in the variables $x_1, \ldots, x_n$. In equation (7.53) let $\alpha = \mu + \delta$ and multiply by p_r. We get

$$a_{\mu+\delta} p_r = \sum_{j=1}^n a_{\mu+\delta+r\epsilon_j}, \tag{7.73}$$

where ϵ_j is the sequence with a 1 in the j-th place and 0 elsewhere. Arrange the sequence $\mu + \delta + r\epsilon_j$ in descending order. If it has two terms equal, then it will contribute nothing to (7.73). Otherwise there is some $p \le q$ for which

$$\mu_{p-1} + n - p + 1 > \mu_q + n - q + r > \mu_p + n - p,$$

in which case $a_{\mu+\delta+r\epsilon_j} = (-1)^{q-p} a_{\lambda+\delta}$, where λ is the partition

$$\lambda = (\mu_1, \ldots, \mu_{p-1}, \mu_q + p - q + r, \mu_p + 1, \ldots, \mu_{q-1} + 1, \mu_{q+1}, \ldots, \mu_n).$$

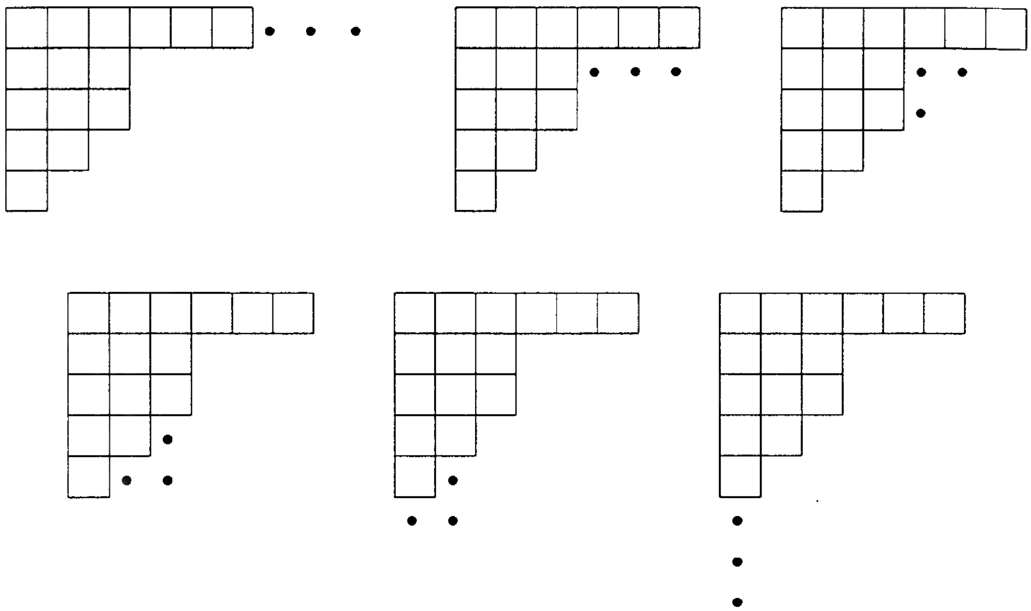


Figure 7-7. Border strips $\lambda/63321$ of size three.

Such partitions are precisely those for which λ/μ is a border strip B of size r, and $q - p$ is just $\mathrm{ht}(B)$. Hence

$$a_{\mu+\delta} p_r = \sum_{\lambda} (-1)^{\mathrm{ht}(\lambda/\mu)} a_{\lambda+\delta}.$$

Divide by a_δ and let $n \to \infty$ to obtain (7.72). □

7.17.2 Example. (a) Let $\mu = 63321$. The border strips of size 3 that can be added to μ are shown in Figure 7-7. Hence

$$s_{63321} p_3 = s_{93321} + s_{66321} - s_{65421} - s_{63333} - s_{633222} + s_{63321111}.$$

(b) Let $\delta = (n-1, n-2, \ldots, 0)$ as above. There are only *two* border strips of size 2 that can be added to δ, and we get

$$s_\delta p_2 = s_{n+1,n-2,n-3,\ldots,1} - s_{n-1,n-2,\ldots,2,1,1,1}.$$

Let $\alpha = (\alpha_1, \alpha_2, \ldots)$ be a weak composition of n. Define a *border-strip tableau* (or *rim-hook tableau*) of shape λ/μ (where $|\lambda/\mu| = n$) and type α to be an assignment of positive integers to the squares of λ/μ such that

(a) every row and column is weakly increasing,
(b) the integer i appears α_i times, and
(c) the set of squares occupied by i forms a border strip.

Equivalently, one may think of a border-strip tableau as a sequence $\mu = \lambda^0 \subseteq \lambda^1 \subseteq \cdots \subseteq \lambda^r \subseteq \lambda$ of partitions such that each skew shape λ^i/λ^{i+1} is a border-strip of

size α_i (including the empty border-strip $\emptyset$ when $\alpha_i = 0$). For instance, the array

$$\begin{array}{ccccccc} 1 & 1 & 1 & 1 & 6 & 6 & 6 \\ 1 & 2 & 2 & 5 & 6 & & \\ 3 & 3 & 5 & 5 & 6 & & \\ 3 & 5 & 5 & 6 & 6 & & \end{array}$$

is a border-strip tableau of shape 7555 and type $(5, 2, 3, 0, 5, 7)$. (The border-strip outlines have been drawn in for the sake of clarity.) Define the *height* $\mathrm{ht}(T)$ of a border-strip tableau T to be

$$\mathrm{ht}(T) = \mathrm{ht}(B_1) + \mathrm{ht}(B_2) + \cdots + \mathrm{ht}(B_k),$$

where $B_1, \ldots, B_k$ are the (nonempty) border strips appearing in T. For the example above we have $\mathrm{ht}(T) = 1 + 0 + 1 + 2 + 3 = 7$.

If we iterate Theorem 7.17.1, successively multiplying s_μ by $p_{\alpha_1}, p_{\alpha_2}, \ldots$, then we obtain immediately the following result.

7.17.3 Theorem. *We have*

$$s_\mu p_\alpha = \sum_\lambda \chi^{\lambda/\mu}(\alpha) s_\lambda, \tag{7.74}$$

where

$$\chi^{\lambda/\mu}(\alpha) = \sum_T (-1)^{\mathrm{ht}(T)}, \tag{7.75}$$

summed over all border-strip tableaux of shape λ/μ and type α.

Taking $\mu = \emptyset$ in Theorem 7.17.3 yields:

7.17.4 Corollary. *We have*

$$p_\alpha = \sum_\lambda \chi^\lambda(\alpha) s_\lambda, \tag{7.76}$$

where $\chi^\lambda(\alpha)$ is given by (7.75).

If we restrict ourselves to n variables where $n \geq \ell(\lambda)$ and apply Theorem 7.15.1, then equation (7.76) may be rewritten

$$p_\alpha a_\delta = \sum_\lambda \chi^\lambda(\alpha) a_{\lambda+\delta}.$$

Hence we obtain the following "formula" for $\chi^\lambda(\alpha)$:

$$\chi^\lambda(\alpha) = [x^{\lambda+\delta}] p_\alpha a_\delta. \tag{7.77}$$

It is easy to use equation (7.74) to express $s_{\lambda/\mu}$ in terms of the power sums. This result (at least in the case $\mu = \emptyset$) is known as the *Murnaghan–Nakayama rule*.

7.17.5 Corollary. *We have*

$$s_{\lambda/\mu} = \sum_{\nu} z_\nu^{-1} \chi^{\lambda/\mu}(\nu) p_\nu, \tag{7.78}$$

where $\chi^{\lambda/\mu}(\nu)$ *is given by (7.75).*

Proof. We have from (7.74) that

$$\begin{aligned}\chi^{\lambda/\mu}(\nu) &= \langle s_\mu p_\nu, s_\lambda \rangle \\ &= \langle p_\nu, s_{\lambda/\mu} \rangle,\end{aligned}$$

and the proof follows from Proposition 7.9.3. □

The orthogonality properties of the bases $\{s_\lambda\}$ and $\{p_\lambda\}$ translate into orthogonality relations satisfied by the coefficients $\chi_\lambda(\mu)$.

7.17.6 Proposition. (a) *Fix* μ, ν. *Then*

$$\sum_\lambda \chi^\lambda(\mu)\chi^\lambda(\nu) = z_\mu \delta_{\mu\nu}.$$

(b) *Fix* λ, μ. *Then*

$$\sum_\nu z_\nu^{-1} \chi^\lambda(\nu)\chi^\mu(\nu) = \delta_{\lambda\mu}.$$

Proof. (a) Expand p_μ and p_ν by (7.76) and take $\langle p_\mu, p_\nu \rangle$.
(b) Expand s_λ and s_μ by (7.78) and take $\langle s_\lambda, s_\mu \rangle$. □

Proposition 7.17.6 is equivalent to the statement that the matrix $(\chi^\lambda(\mu) z_\mu^{-1/2})_{\lambda,\mu \vdash n}$ is an orthogonal matrix. This may be seen directly from the fact that this matrix is the transition matrix between the two orthonormal bases $\{s_\lambda\}$ and $\{p_\mu z_\mu^{-1/2}\}$.

A remarkable consequence of Corollary 7.17.4 is that the coefficients $\chi^\lambda(\alpha)$ do not depend on the order of the entries of α (since the same is true of the product $p_\alpha = p_{\alpha_1} p_{\alpha_2} \cdots$). This fact can be of great value in obtaining information about the numbers $\chi^\lambda(\alpha)$. As a sample application, we mention the following result.

7.17.7 Proposition. *Let* δ *be the "staircase shape"* $\delta = (m-1, m-2, \ldots, 1)$. *Then* s_δ *is a polynomial in the odd power sums* $p_1, p_3, \ldots$.

Proof. We need to show that $\chi^\delta(\nu) = 0$ if ν has an even part. Let α be an ordering of the parts of ν such that the *last* nonzero entry α_k of α is even. Thus the border-strip tableaux T in (7.75) have the property that the squares labeled k form a border strip δ/ν of size α_k. But every border strip δ/ν has odd size, so no such T exists. □

For the converse to Proposition 7.17.7, see Exercise 7.54.

Note (for algebraists). The coefficients $\chi^\lambda(\nu)$ for $\lambda, \nu \vdash n$ have a fundamental algebraic interpretation: They are the values of the irreducible (ordinary) characters of the symmetric group $\mathfrak{S}_n$. More precisely, the irreducible characters χ^λ of $\mathfrak{S}_n$ are indexed in a natural way by partitions $\lambda \vdash n$, and $\chi^\lambda(\nu)$ is the value of χ^λ at an element $w \in \mathfrak{S}_n$ of cycle type ν. Thus Proposition 7.17.6 is just the standard orthogonality relations satisfied by irreducible characters. Now it may be seen e.g. immediately from (7.75) that the degree (or dimension) of the character χ^λ is given by

$$\deg \chi^\lambda := \chi^\lambda(1^n) = f^\lambda. \tag{7.79}$$

Thus Corollary 7.12.6 agrees with the well-known result that for any finite group G,

$$\sum_{\chi \in \hat{G}} (\dim \chi)^2 = \#G, \tag{7.80}$$

where $\hat{G}$ is the set of irreducible characters of G. Moreover, Corollary 7.13.9 agrees with the less well-known result that

$$\sum_{\chi \in \hat{G}} \dim \chi = \#\{w \in G : w^2 = 1\}$$

if and only if every (ordinary) representation of G is equivalent to a real representation. For further information on the connections between symmetric functions and the characters of $\mathfrak{S}_n$, see the next section and many of the exercises for this chapter.

7.18 The Characters of the Symmetric Group

This section is not needed for the rest of the text (with a few minor exceptions) and assumes a basic knowledge of the representation theory of finite groups. Our goal will be to show that the functions χ^λ of the previous section (where $\chi^\lambda(\mu)$ is interpreted as $\chi^\lambda(w)$ when w is an element of $\mathfrak{S}_n$ of (cycle) type μ) are the irreducible characters of $\mathfrak{S}_n$.

Let CF^n denote the set of all class functions (i.e., functions constant on conjugacy classes) $f : \mathfrak{S}_n \to \mathbb{Q}$. Recall that CF^n has a natural scalar product defined by

$$\langle f, g \rangle = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} f(w) g(w).$$

Sometimes by abuse of notation we write $\langle \phi, \gamma \rangle$ instead of $\langle f, g \rangle$ when ϕ and γ are representations of $\mathfrak{S}_n$ with characters f and g.

Note. For general finite groups G, we can define $\mathrm{CF}(G)$ to be the set of all class functions $f : G \to \mathbb{C}$, and we can define the scalar product on $\mathrm{CF}(G)$ by

$$\langle f, g \rangle = \frac{1}{\#G} \sum_{w \in G} f(w) \bar{g}(w),$$

where $\bar{g}(w)$ denotes the complex conjugate of $g(w)$. Since all (complex) characters of $\mathfrak{S}_n$ turn out to be rational, it suffices to use the ground field $\mathbb{Q}$ instead of $\mathbb{C}$ when dealing with the characters of $\mathfrak{S}_n$.

Now let us recall some basic facts from the theory of permutation representations. If X is a finite set and G a finite group, then an *action* of G on X is a homomorphism $\varphi : G \to \mathfrak{S}_X$. If $s \in X$ and $w \in G$, then we write $w \cdot s$ for $\varphi(w)(s)$. The action of G on X extends to an action on $\mathbb{C}X$ (the complex vector space with basis X) by linearity. Hence φ can be regarded as a linear representation $\varphi : G \to \mathrm{GL}(\mathbb{C}X)$. The character of this representation is given by

$$\chi^{\varphi}(w) = \operatorname{tr} \varphi(w) = \#\mathrm{Fix}(w),$$

where $\mathrm{Fix}(w) = \{s \in X \,:\, w \cdot s = s\}$, the set of points fixed by w.

The action $\varphi : G \to \mathfrak{S}_X$ is *transitive* if for any $s, t \in X$ there is a $w \in G$ satisfying $w \cdot s = t$. If H is a subgroup of G, then G acts on the set G/H of left cosets of G by $w \cdot vH = wvH$. (We do not assume H is a normal subgroup, so G/H need not have the structure of a group.) Every transitive action of G is equivalent to an action on the left cosets of some subgroup H. Moreover, this action is equivalent to $\mathrm{ind}_H^G 1_H$, the *induction* from H to G of the trivial representation 1_H of H. We sometimes abbreviate this representation as 1_H^G. The well-known "Burnside's lemma" (see Lemma 7.24.5) is equivalent to the statement that

$$\langle 1_H^G, 1_G \rangle = \#\text{ of orbits of } G \text{ acting on } G/H. \tag{7.81}$$

Here $\langle 1_H^G, 1_G \rangle$ denotes the multiplicity of the trivial representation 1_G of G in 1_H^G, given more explicitly by

$$\langle 1_H^G, 1_G \rangle = \frac{1}{\#G} \sum_{w \in G} \#\mathrm{Fix}(w).$$

In the above sum $\mathrm{Fix}(w)$ refers to the action of G on the set G/H, so that $\mathrm{Fix}(w)$ is just the value of the character of this action on w.

Our present goal is to find "enough" subgroups H of $\mathfrak{S}_n$ so that we can obtain all the irreducible characters of $\mathfrak{S}_n$ as linear combinations of characters of the representations $1_H^{\mathfrak{S}_n}$. To this end, if $\alpha = (\alpha_1, \dots, \alpha_\ell) \in \mathbb{P}^\ell$ and $|\alpha| := \alpha_1 + \cdots + \alpha_\ell = n$, then define the *Young subgroup* $\mathfrak{S}_\alpha \subseteq \mathfrak{S}_n$ to be

$$\mathfrak{S}_\alpha = \mathfrak{S}_{\alpha_1} \times \mathfrak{S}_{\alpha_2} \times \cdots \times \mathfrak{S}_{\alpha_\ell},$$

where $\mathfrak{S}_{\alpha_1}$ permutes $1, 2, \dots, \alpha_1$; $\mathfrak{S}_{\alpha_2}$ permutes $\alpha_1 + 1, \alpha_1 + 2, \dots, \alpha_1 + \alpha_2$, etc. If α and β differ from each other only by a permutation of coordinates, then $\mathfrak{S}_\alpha$ and $\mathfrak{S}_\beta$ are conjugate subgroups of $\mathfrak{S}_n$, and the representations $1_{\mathfrak{S}_\alpha}^{\mathfrak{S}_n}$ and $1_{\mathfrak{S}_\beta}^{\mathfrak{S}_n}$ are equivalent and hence have the same character. In particular, there is a unique $\lambda \vdash n$ for which $\mathfrak{S}_\alpha$ and $\mathfrak{S}_\lambda$ are conjugate.

It is important to understand the combinatorial significance of the representations $1_{\mathfrak{S}_\alpha}^{\mathfrak{S}_n}$, so we will explain this topic in some detail. If we write a permutation w in the usual way as $w_1 w_2 \cdots w_n$ (so w_i is the *value* $w(i)$ of w at i), then it is easy to see that every left coset $v\mathfrak{S}_\alpha$ contains a unique permutation that is an α*-shuffle*, i.e., $1, 2, \dots, \alpha_1$ appear in increasing order, $\alpha_1 + 1, \alpha_1 + 2, \dots, \alpha_1 + \alpha_2$ appear in increasing order, etc. For instance, one of the left cosets of $\mathfrak{S}_{(2,2)}$ is given by $\{1324, 1423, 2314, 2413\}$, which contains the unique α-shuffle 1324. We can identify an α-shuffle with a permutation of the multiset $M_\alpha = \{a^{\alpha_1}, b^{\alpha_2}, \dots\}$, by replacing $1, 2, \dots, \alpha_1$ with a; $\alpha_1+1, \alpha_1+2, \dots, \alpha_1+\alpha_2$ with b; etc. $\mathfrak{S}_n$ then acts on a permutation π of M_α by permuting *positions*. For instance, $2431 \cdot baab = abab$, since the second element of $baab$ is moved to the first position, the fourth element to the second position, etc.

Alternatively, we can write a permutation $w \in \mathfrak{S}_n$ as the word $w^{-1}(1)w^{-1}(2)\cdots w^{-1}(n)$, so $w^{-1}(i)$ is the *position* of i in the word $w_1 w_2 \cdots w_n$. With this representation of permutations, every left coset of $\mathfrak{S}_\alpha$ contains a unique word $w' = w'_1 w'_2 \cdots w'_n$ such that $w'_1 < w'_2 < \cdots < w'_{\alpha_1}$, $w'_{\alpha_1+1} < w'_{\alpha_1+2} < \cdots < w'_{\alpha_1+\alpha_2}$, etc. Equivalently, the descent set $D(w')$ is contained in the set $S_\alpha = \{\alpha_1, \alpha_1 + \alpha_2, \dots, \alpha_1 + \cdots + \alpha_{\ell-1}\}$. We can also view these distinguished coset representatives as α*-flags*, i.e., chains $\emptyset = N_0 \subset N_1 \subset \cdots \subset N_\ell = [n]$ of subsets of $[n]$ such that $\#(N_i - N_{i-1}) = \alpha_i$, viz., $N_i = \{w'_1, w'_2, \dots, w'_{\alpha_1+\alpha_2+\cdots+\alpha_i}\}$. $\mathfrak{S}_n$ then acts on a flag F by permuting *elements*. For instance, if F is given by $\emptyset \subset 24 \subset 245 \subset 123456$ (so that $\alpha = (2, 1, 3)$) and if $w = 523614$, then $w \cdot F$ is given by $\emptyset \subset 26 \subset 126 \subset 123456$, since 2 and 6 are in the second and fourth position of w, 1 is in the fifth position of w, and 3, 4, 5 are in third, sixth, and fourth position of w.

Some special cases of the action of $\mathfrak{S}_n$ on α-flags should be noted. If $\alpha = (k, n-k)$, then an α-flag $\emptyset \subset N \subset [n]$ is equivalent to the k-subset N of $[n]$, and the action of $w \in \mathfrak{S}_n$ on F is equivalent to the "standard" action of $\mathfrak{S}_n$ on N that replaces $i \in N$ with $w^{-1}(i)$. We may write this equivalence as

$$\mathfrak{S}_n/(\mathfrak{S}_k \times \mathfrak{S}_{n-k}) \cong \binom{[n]}{k}. \tag{7.82}$$

Similarly, if $\alpha = (1, 1, n-2)$ then we can identify the α-flag $\emptyset \subset \{a\} \subset \{a, b\} \subset [n]$ with the ordered pair (a, b) of distinct elements of $[n]$, so we can write

$$\mathfrak{S}_n/(\mathfrak{S}_1 \times \mathfrak{S}_1 \times \mathfrak{S}_{n-2}) \cong [n] \times [n] - \{(a, a) : a \in [n]\}. \tag{7.83}$$

Similar interpretations can be given for various other values of α.

Our main tool will be a linear transformation $\mathrm{ch} : \mathrm{CF}^n \to \Lambda^n$ called the *(Frobenius) characteristic map*. If $f \in \mathrm{CF}^n$, then define

$$\begin{aligned} \mathrm{ch}\, f &= \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} f(w) p_{\rho(w)} \\ &= \sum_\mu z_\mu^{-1} f(\mu) p_\mu, \end{aligned}$$

where $f(\mu)$ denotes $f(w)$ for any w of type $\rho(w) = \mu$. Equivalently, extending the ground field $\mathbb{Q}$ to the ring Λ and defining $\Psi(w) = p_{\rho(w)}$, we have

$$\operatorname{ch} f = \langle f, \Psi \rangle. \tag{7.84}$$

Note that if f_μ is the class function defined by

$$f_\mu(w) = \begin{cases} 1 & \text{if } \rho(w) = \mu \\ 0 & \text{otherwise,} \end{cases}$$

then $\operatorname{ch} f_\mu = z_\mu^{-1} p_\mu$.

NOTE. Let $\varphi : \mathfrak{S}_n \to \mathrm{GL}(V)$ be a representation of $\mathfrak{S}_n$ with character χ. Sometimes by abuse of notation we will write $\operatorname{ch}\varphi$ or $\operatorname{ch} V$ instead of $\operatorname{ch}\chi$. We also sometimes call the symmetric function $\operatorname{ch}\chi$ ($= \operatorname{ch}\varphi = \operatorname{ch} V$) the *Frobenius image* of χ, φ, or V.

7.18.1 Proposition. *The linear transformation* ch *is an isometry, i.e.,*

$$\langle f, g \rangle_{\mathrm{CF}^n} = \langle \operatorname{ch} f, \operatorname{ch} g \rangle_{\Lambda^n}.$$

Proof. We have (using Proposition 7.9.3)

$$\begin{aligned} \langle \operatorname{ch} f, \operatorname{ch} g \rangle &= \left\langle \sum_\mu z_\lambda^{-1} f(\lambda) p_\mu, \sum_\mu z_\mu^{-1} g(\mu) p_\mu \right\rangle \\ &= \sum_\lambda z_\lambda^{-1} f(\lambda) g(\lambda) \\ &= \langle f, g \rangle. \end{aligned}$$

□

We now want to define a product on class functions that will correspond to the ordinary product of symmetric functions under the characteristic map ch. Let $f \in \mathrm{CF}^m$ and $g \in \mathrm{CF}^n$. Define the (pointwise) product $f \times g \in \mathrm{CF}(\mathfrak{S}_m \times \mathfrak{S}_n)$ by

$$(f \times g)(u, v) = f(u) g(v).$$

If f and g are characters of representations φ and ψ, then $f \times g$ is just the character of the tensor product representation $\varphi \otimes \psi$ of $\mathfrak{S}_m \times \mathfrak{S}_n$. Now define the *induction product* $f \circ g$ of f and g to be the induction of $f \times g$ to $\mathfrak{S}_{m+n}$, where as before $\mathfrak{S}_m$ permutes $1, 2, \ldots, m$ while $\mathfrak{S}_n$ permutes $m+1, m+2, \ldots, m+n$. In symbols.

$$f \circ g = \operatorname{ind}_{\mathfrak{S}_m \times \mathfrak{S}_n}^{\mathfrak{S}_{m+n}} (f \times g).$$

Let $\mathrm{CF} = \mathrm{CF}^0 \oplus \mathrm{CF}^1 \oplus \cdots$, and extend the scalar product on CF^n to all of CF by setting $\langle f, g \rangle = 0$ if $f \in \mathrm{CF}^m$, $g \in \mathrm{CF}^n$, and $m \neq n$. The induction product

on characters extends to all of CF by (bi)linearity. It is not hard to check that this makes CF into an associative commutative graded $\mathbb{Q}$-algebra with identity $1 \in \mathrm{CF}^0$. Similarly we can extend the characteristic map ch to a linear transformation $\mathrm{ch} : \mathrm{CF} \to \Lambda$.

7.18.2 Proposition. *The characteristic map* $\mathrm{ch} : \mathrm{CF} \to \Lambda$ *is a bijective ring homomorphism, i.e.,* ch *is one-to-one and onto, and satisfies*

$$\mathrm{ch}(f \circ g) = (\mathrm{ch}\, f)(\mathrm{ch}\, g).$$

Proof. Let $\mathrm{res}_H^G f$ denote the restriction of the class function f on G to the subgroup H. We then have

$$\begin{aligned}
\mathrm{ch}(f \circ g) &= \mathrm{ch}\big(\mathrm{ind}_{\mathfrak{S}_m \times \mathfrak{S}_n}^{\mathfrak{S}_{m+n}} (f \times g)\big) \\
&= \big\langle \mathrm{ind}_{\mathfrak{S}_m \times \mathfrak{S}_n}^{\mathfrak{S}_{m+n}} (f \times g), \Psi \big\rangle \qquad \text{(by (7.84))} \\
&= \big\langle f \times g, \mathrm{res}_{\mathfrak{S}_m \times \mathfrak{S}_n}^{\mathfrak{S}_{m+n}} \Psi \big\rangle_{\mathfrak{S}_m \times \mathfrak{S}_n} \qquad \text{(by Frobenius reciprocity)} \\
&= \frac{1}{m!\,n!} \sum_{u \in \mathfrak{S}_m} \sum_{v \in \mathfrak{S}_n} f(u) g(v) \Psi(uv) \\
&= \frac{1}{m!\,n!} \sum_{u \in \mathfrak{S}_m} \sum_{v \in \mathfrak{S}_n} f(u) g(v) \Psi(u) \Psi(v) \\
&= \langle f, \Psi \rangle_{\mathfrak{S}_m} \langle g, \Psi \rangle_{\mathfrak{S}_n} \\
&= (\mathrm{ch}\, f)(\mathrm{ch}\, g).
\end{aligned}$$

Moreover, from the definition of ch and the fact that the power sums p_μ form a $\mathbb{Q}$-basis for Λ it follows that ch is bijective. □

We wish to apply Proposition 7.18.2 to evaluate the characteristic map at the character η^α of the representation $1_{\mathfrak{S}_\alpha}^{\mathfrak{S}_n}$ discussed above. First note that by equation (7.22) and the definition of ch we have

$$\mathrm{ch}\, 1_{\mathfrak{S}_n} = \sum_{\lambda \vdash n} z_\lambda^{-1} p_\lambda = h_n. \tag{7.85}$$

7.18.3 Corollary. *We have* $\mathrm{ch}\, 1_{\mathfrak{S}_\alpha}^{\mathfrak{S}_n} = h_\alpha$.

Proof. Since $1_{\mathfrak{S}_\alpha}^{\mathfrak{S}_n} = 1_{\mathfrak{S}_{\alpha_1}}^{\mathfrak{S}_n} \circ 1_{\mathfrak{S}_{\alpha_2}}^{\mathfrak{S}_n} \circ \cdots \circ 1_{\mathfrak{S}_{\alpha_\ell}}^{\mathfrak{S}_n}$, the proof follows from Proposition 7.18.2 and equation (7.85). □

Now let R^n denote the set of all *virtual characters* of $\mathfrak{S}_n$, i.e., functions on $\mathfrak{S}_n$ that are the difference of two characters. Thus R^n is a lattice (discrete subgroup of maximum rank) in the vector space CF^n. The rank of R^n is $p(n)$, the number of

partitions of n, and a basis consists of the irreducible characters of $\mathfrak{S}_n$. (Recall that in any finite group G, the number of linearly independent irreducible characters of G is equal to the number of conjugacy classes in G. For $\mathfrak{S}_n$ the number of conjugacy classes is $p(n)$.) This basis is the unique orthonormal basis for R^n up to sign and order, since the transition matrix between two such bases must be a integral orthogonal matrix and hence a signed permutation matrix. (See the note after the proof of Corollary 7.12.2 for similar reasoning.) Define $R = R^0 \oplus R^1 \oplus \cdots$.

7.18.4 Proposition. *The image of R under the characteristic map* ch *is $\Lambda_{\mathbb{Z}}$. Hence* ch $: R \to \Lambda_{\mathbb{Z}}$ *is a ring isomorphism.*

Proof. It will suffice to find integer linear combinations of the characters η^α of the representations $1_{\mathfrak{S}_\alpha}^{\mathfrak{S}_n}$ that are the irreducible characters of $\mathfrak{S}_n$. The Jacobi–Trudi identity (Theorem 7.16.1) suggests that we define the (possibly virtual) characters $\psi^\lambda = \det(\eta^{\lambda_i - i + j})$, where the product used in evaluating the determinant is the induction product. Then by the Jacobi–Trudi identity and Proposition 7.18.2 we have

$$\operatorname{ch}(\psi^\lambda) = s_\lambda. \tag{7.86}$$

Since ch is an isometry (Proposition 7.18.1) we get $\langle \psi^\lambda, \psi^\mu \rangle = \delta_{\lambda\mu}$. As pointed out above, this means that the class functions ψ^λ are, up to sign, the irreducible characters of $\mathfrak{S}_n$. Hence the ψ^λ for $\lambda \vdash n$ form a $\mathbb{Z}$-basis for R^n, and the image of R^n is the $\mathbb{Z}$-span of the s_λ's, which is just Λ^n as claimed. □

Finally we come to the main result of this section.

7.18.5 Theorem. *Regard the functions χ^λ (where $\lambda \vdash n$) of Section 7.17 as functions on $\mathfrak{S}_n$ given by $\chi^\lambda(w) = \chi^\lambda(\mu)$, where w has cycle type μ. Then the χ^λ's are the irreducible characters of the symmetric group $\mathfrak{S}_n$.*

Proof. By the Murnaghan–Nakayama rule (Corollary 7.17.5), we have

$$\operatorname{ch}(\chi^\lambda) = \sum_\mu z_\mu^{-1} \chi^\lambda(\mu) p_\mu = s_\lambda.$$

Hence by equation (7.86), we get $\chi^\lambda = \psi^\lambda$. Since the ψ^λ's, up to sign, are the irreducible characters of $\mathfrak{S}_n$, it remains only to determine whether χ^λ or $-\chi^\lambda$ is a character. But $\chi^\lambda(1^n) = f^\lambda > 0$, so χ^λ is an irreducible character. □

Note. We have described a natural way to index the irreducible characters of $\mathfrak{S}_n$ by partitions of n, while the cycle type of a permutation defines a natural indexing of the conjugacy classes of $\mathfrak{S}_n$ by partitions of n. Hence we have a canonical bijection between the conjugacy classes and the irreducible characters of $\mathfrak{S}_n$. However, this bijection is essentially "accidental" and does not have any useful

properties. For arbitrary finite groups there is in general no canonical bijection between irreducible characters and conjugacy classes.

7.18.6 Corollary. *Let $\mu \vdash m$, $\nu \vdash n$, $\lambda \vdash m+n$. Then the Littlewood–Richardson coefficient $c^\lambda_{\mu\nu}$ of equation (7.64) is nonnegative*

Proof. By Propositions 7.18.1 and 7.18.2 we have

$$c^\lambda_{\mu\nu} = \langle s_\lambda, s_\mu s_\nu \rangle = \langle \chi^\lambda, \chi^\mu \circ \chi^\nu \rangle.$$

Since by Theorem 7.18.5 χ^μ and χ^ν are characters of $\mathfrak{S}_m$ and $\mathfrak{S}_n$, respectively, it follows from the basic theory of induced characters that $\chi^\mu \circ \chi^\nu$ is a character of $\mathfrak{S}_{m+n}$. Hence $\langle \chi^\lambda, \chi^\mu \circ \chi^\nu \rangle \geq 0$. □

Combinatorial descriptions of the numbers $c^\nu_{\lambda\mu}$ are given in Appendix 1, Section A1.3. A primary use of Theorem 7.18.5 is the following. Let χ be any character (or even a virtual character) of $\mathfrak{S}_n$. Theorem 7.18.5 shows that the problem of decomposing χ into irreducibles is equivalent to expanding $\operatorname{ch}\chi$ into Schur functions. Thus all the symmetric function machinery we have developed can be brought to bear on the problem of decomposing χ. An important example is given by the characters η^α of the representations $1_{\mathfrak{S}_\alpha}^{\mathfrak{S}_n}$. The result that expresses η^α in terms of irreducibles is known as *Young's rule*.

7.18.7 Proposition. *Let α be a composition of n and $\lambda \vdash n$. Then the multiplicity of the irreducible character χ^λ in the character η^α is just the Kostka number $K_{\lambda\alpha}$. In symbols,*

$$\langle \eta^\alpha, \chi^\lambda \rangle = K_{\lambda\alpha}.$$

Proof. By Corollary 7.18.3 we have $\operatorname{ch}\eta^\alpha = h_\alpha$. The proof then follows from Corollary 7.12.4 and Theorem 7.18.5. □

7.18.8 Example. (a) Let χ denote the character of the action of $\mathfrak{S}_n$ on the k-subsets of $[n]$ (by permuting elements of $[n]$). By equation (7.82), we have $\operatorname{ch}\chi = h_k h_{n-k}$. Assume without loss of generality that $k \leq n/2$ (since the action on k-element subsets is equivalent to the action on $(n-k)$-element subsets). The multiplicity of χ^λ in χ is the number of SSYTs of shape λ and type $1^{n-k}2^k$. There is one such SSYT if $\lambda = (n-m, m)$ with $0 \leq m \leq k$, and no SSYTs otherwise. Hence

$$\chi = \chi^{(n)} + \chi^{(n-1,1)} + \cdots + \chi^{(n-k,k)}.$$

Note the special case $k = 1$; this corresponds to the "defining representation" of $\mathfrak{S}_n$ (the action of $\mathfrak{S}_n$ on $[n]$), with character $\chi^{(n)} + \chi^{(n-1,1)}$.

(b) Let χ denote the character of the action of $\mathfrak{S}_n$ on ordered pairs (i, j) of distinct elements of $[n]$. By equation (7.83), we have $\operatorname{ch}\chi = h_1^2 h_{n-2}$. There are five kinds of SSYT of type $(n-2, 1, 1)$, viz.,

$$\begin{matrix} 1\cdots 123 & 1\cdots 12 & 1\cdots 13 & 11\cdots 1 & 1\cdots 1 \\ & 3 & 2 & 23 & 2 \\ & & & & 3 \end{matrix}$$

Hence

$$\chi = \chi^{(n)} + 2\chi^{(n-1,1)} + \chi^{(n-2,2)} + \chi^{(n-2,1,1)}.$$

(c) The *regular representation* of $\mathfrak{S}_n$ is defined to be the action of $\mathfrak{S}_n$ on itself by left multiplication. Hence it is given by $1_{\mathfrak{S}_1\times\cdots\times\mathfrak{S}_1}^{\mathfrak{S}_n}$, whose Frobenius image is h_1^n (by Corollary 7.18.3). By Corollary 7.12.5 we have $h_1^n = \sum_{\lambda\vdash n} f^\lambda s_\lambda$. Hence the multiplicity in the regular representation of the irreducible representation of $\mathfrak{S}_n$ whose character is χ^λ is just $f^\lambda = \chi^\lambda(1^n)$. Thus Corollary 7.12.5 is a symmetric function statement for $\mathfrak{S}_n$ of the fact that the multiplicity of an irreducible representation of a finite group in the regular representation is equal to its degree.

7.19 Quasisymmetric Functions

We have succeeded in expanding the Schur functions in terms of the four bases m_λ, h_λ, e_λ, and p_λ. We have also given a formula for s_λ (in n variables) as a quotient of determinants. There is one further expression for s_λ which has many combinatorial ramifications. We will write s_λ in terms of a basis not of the space Λ, but of a larger space $\mathcal{Q}$. The theory of P-partitions, as discussed in Section 4.5, will be generalized in order to obtain this expansion of s_λ (and more generally of $s_{\lambda/\mu}$).

A *quasisymmetric function* in the variables $x_1, x_2, \ldots$, say with rational coefficients, is a formal power series $f = f(x) \in \mathbb{Q}[[x_1, x_2, \ldots]]$ of bounded degree such that for any $a_1, \ldots, a_k \in \mathbb{P}$ we have

$$\left[x_{i_1}^{a_1}\cdots x_{i_k}^{a_k}\right] f = \left[x_{j_1}^{a_1}\cdots x_{j_k}^{a_k}\right] f$$

whenever $i_1 < \cdots < i_k$ and $j_1 < \cdots < j_k$. Clearly every symmetric function is quasisymmetric, but not conversely. For instance, the series $\sum_{i<j} x_i^2 x_j$ is quasisymmetric but not symmetric.

Let $\mathcal{Q}^n$ denote the set of all homogeneous quasisymmetric functions of degree n. It is clear that $\mathcal{Q}^n$ is a vector space (over $\mathbb{Q}$). We will be indexing certain elements of $\mathcal{Q}^n$ by compositions $\alpha = (\alpha_1, \ldots, \alpha_k)$ of n. We will often be using the natural one-to-one correspondence between compositions α of n and subsets S of $[n-1]$. Thus we associate the set $S_\alpha = \{\alpha_1, \alpha_1+\alpha_2, \ldots, \alpha_1+\cdots+\alpha_{k-1}\}$ with the composition α, and the composition $\operatorname{co}(S) = (s_1, s_2 - s_1, s_3 - s_2, \ldots, n - s_{k-1})$ with the set $S = \{s_1, s_2, \ldots, s_{k-1}\}_<$. It is clear that $\operatorname{co}(S_\alpha) = \alpha$ and $S_{\operatorname{co}(S)} = S$. We extend the definition of co to permutations $w \in \mathfrak{S}_n$ by defining $\operatorname{co}(w) = \operatorname{co}(D(w))$, where $D(w)$ denotes the descent set of w.

Let Comp(n) denote the set of compositions of n, so #Comp(n) $= 2^{n-1}$. Given $\alpha \in$ Comp(n), define the *monomial quasisymmetric function* M_α by

$$M_\alpha = \sum_{i_1 < \cdots < i_k} x_{i_1}^{\alpha_1} \cdots x_{i_k}^{\alpha_k}. \tag{7.87}$$

If $f \in \mathcal{Q}^n$ then it is clear that

$$f = \sum_{\alpha \in \mathrm{Comp}(n)} \left([x_1^{\alpha_1} \cdots x_k^{\alpha_k}] f \right) M_\alpha. \tag{7.88}$$

From (7.88) it follows that the set $\{M_\alpha : \alpha \in \mathrm{Comp}(n)\}$ is a basis for $\mathcal{Q}^n$. In particular,

$$\dim \mathcal{Q}^n = 2^{n-1}.$$

It is an easy exercise to see that if $f \in \mathcal{Q}^m$ and $g \in \mathcal{Q}^n$, then $fg \in \mathcal{Q}^{m+n}$. (See Exercise 7.93 for a more precise result.) Hence if $\mathcal{Q} = \mathcal{Q}^0 \oplus \mathcal{Q}^1 \oplus \cdots$, then $\mathcal{Q}$ is a $\mathbb{Q}$-algebra, called the *algebra* (or *ring*) *of quasisymmetric functions* (over $\mathbb{Q}$).

We will now consider another important basis for $\mathcal{Q}^n$. Given $\alpha \in$ Comp(n), define

$$L_\alpha = \sum_{\substack{i_1 \le \cdots \le i_n \\ i_j < i_{j+1} \text{ if } j \in S_\alpha}} x_{i_1} x_{i_2} \cdots x_{i_n}. \tag{7.89}$$

It is clear that $L_\alpha \in \mathcal{Q}^n$. We call L_α a *fundamental quasisymmetric function.*

7.19.1 Proposition. *For $\alpha \in$ Comp(n) we have*

$$L_\alpha = \sum_{S_\alpha \subseteq T \subseteq [n-1]} M_{\mathrm{co}(T)} \tag{7.90}$$

$$M_\alpha = \sum_{S_\alpha \subseteq T \subseteq [n-1]} (-1)^{\#(T - S_\alpha)} L_{\mathrm{co}(T)}. \tag{7.91}$$

Hence the set $\{L_\alpha : \alpha \in \mathrm{Comp}(n)\}$ is a basis for $\mathcal{Q}^n$.

Proof. Equation (7.90) is an immediate consequence of (7.89), on grouping the sequences $i_1 \le \cdots \le i_n$ according to whether $i_j < i_{j+1}$ or $i_j = i_{j+1}$ for each j. Equation (7.91) then follows from (7.90) by the Principle of Inclusion–Exclusion (Theorem 2.1.1). □

It is natural to ask under what conditions is a quasisymmetric function actually symmetric.

7.19.2 Proposition. *Let $f \in \mathcal{Q}^n$, say $f = \sum_{\alpha \in \mathrm{Comp}(n)} c_\alpha M_\alpha$. Then $f \in \Lambda^n$ if and only if for any two compositions α and β of n that have the same multiset of parts, we have $c_\alpha = c_\beta$.*

Proof. Let $\mathcal{R}^n$ be the subspace of $\mathcal{Q}^n$ consisting of all $f = \sum c_\alpha M_\alpha$ satisfying the conditions of the proposition. Given $\lambda \vdash n$, define $R_\lambda = \sum_\alpha M_\alpha$, summed over all distinct permutations of the parts of λ. It is clear that $\{R_\lambda : \lambda \vdash n\}$ is a basis for $\mathcal{R}^n$, so $\dim \mathcal{R}^n = p(n)$, the number of partitions of n. On the other hand, it is also evident that $R_\lambda = m_\lambda \in \Lambda^n$, so $\mathcal{R}^n \subseteq \Lambda^n$. Since $\dim \Lambda^n = p(n)$, the proof follows. □

Let us now consider the connection between quasisymmetric functions and the theory of P-partitions. If X is any finite set and $f : X \to \mathbb{P}$, then define

$$x^f = \prod_{t \in X} x_{f(t)} = \prod_{i \geq 1} x_i^{\# f^{-1}(i)}.$$

In Definition 4.5.1 we defined the notion of a π-compatible function $f : [n] \to C$, where $\pi = \pi_1 \cdots \pi_n \in \mathfrak{S}_n$ and C is a chain. It is more convenient here to work with the "reverse" notion. We say that f is *reverse π-compatible* if

$$\begin{aligned} &f(\pi_1) \leq f(\pi_2) \leq \cdots \leq f(\pi_n) \\ &f(\pi_i) < f(\pi_{i+1}) \qquad \text{if } \pi_i > \pi_{i+1}. \end{aligned}$$

Clearly Lemma 4.5.1 carries over with "compatible" replaced with "reverse compatible." Thus every $f : [n] \to C$ is reverse π-compatible for a *unique* $\pi \in \mathfrak{S}_n$.

7.19.3 Lemma. *Let $\pi \in \mathfrak{S}_n$, and let S_π^r denote the set of all reverse π-compatible functions $f : [n] \to \mathbb{N}$. Then*

$$\sum_{f \in S_\pi^r} x^f = L_{\mathrm{co}(\pi)}(x).$$

Proof. Immediate from a comparison of the definition (7.89) of L_α and the definition of reverse π-compatible. □

Now let P be a finite poset of cardinality n. To be consistent with our treatment of SSYTs, we will deal with functions $\sigma : P \to \mathbb{P}$ rather than $\sigma : P \to \mathbb{N}$ as in Section 4.5. It is a trivial matter to modify the theory to allow also $\sigma(t) = 0$. A *reverse P-partition* is an order-preserving map $\sigma : P \to \mathbb{P}$ (and so is equivalent to an ordinary P^*-partition, where P^* denotes the dual of P). Let $\mathcal{A}^r(P)$ denote the set of reverse P-partitions. Define

$$K_P(x) = \sum_{\sigma \in \mathcal{A}^r(P)} x^\sigma. \tag{7.92}$$

Note that $K_P(x)$ is a quasisymmetric function that tells us for each weak composition $\alpha = (\alpha_1, \alpha_2, \ldots)$ of n the number of reverse P-partitions with α_i parts equal to i. As in Section 4.5, regard P as a natural partial order on $[n]$, and let

$\mathcal{L}(P) \subseteq \mathfrak{S}_n$ be the Jordan–Hölder set of P. The fundamental decomposition $\mathcal{A}(P) = \dot{\bigcup}_{\pi\in\mathcal{L}(P)} S_\pi$ clearly works just as well in the reverse situation:

$$\mathcal{A}^r(P) = \dot{\bigcup_{\pi\in\mathcal{L}(P)}} S^r_\pi. \tag{7.93}$$

Hence, letting $J(P)$ denote the lattice of order ideals of P as in Chapter 3, we have

$$\begin{aligned} K_P(x) &= \sum_{\pi\in\mathcal{L}(P)} L_{\mathrm{co}(\pi)}(x) \\ &= \sum_{S\subseteq[n-1]} \beta(J(P), S) L_{\alpha_S}(x), \end{aligned} \tag{7.94}$$

the latter equality by Theorem 3.12.1. This result shows that the expansion of a quasisymmetric function in terms of the L_α's can be of combinatorial significance, and that the coefficients may be regarded as an analogue of the numbers $\beta(J(P), S)$.

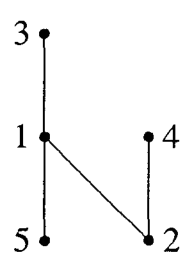
To apply the kind of reasoning of the previous paragraph to Schur functions, we need an extension of the theory of P-partitions to the case where P need not be a natural partial order. Define a *labeling* of P to be a bijection $\omega : P \to [n]$. (Do not confuse the labeling ω with the involution $\omega : \Lambda \to \Lambda$.) Alternatively, one could think of P as a partial ordering on $[n]$ by identifying $t \in P$ with $\omega(t)$. It is convenient, however, to work with labelings and so avoid having to deal with two different orderings on $[n]$.

A *reverse* (P, ω)*-partition* is a map $\sigma : P \to \mathbb{N}$ satisfying the two axioms:

(R1) If $s \le t$ in P then $\sigma(s) \le \sigma(t)$. In other words, σ is *order-preserving*.
(R2) If $s < t$ in P and $\omega(s) > \omega(t)$, then $\sigma(s) < \sigma(t)$.

Thus a reverse (P, ω)-partition is just a reverse P-partition with additional conditions specified by ω as to when *strict* inequality $\sigma(s) < \sigma(t)$ must occur. If, for instance, ω is order-preserving, then a reverse (P, ω)-partition is just a reverse P-partition. On the other hand, if ω is order-reversing, then a reverse (P, ω)-partition is just a strict reverse P-partition.

Let $\mathcal{L}(P, \omega)$ denote the set of all permutations $\pi = \pi_1 \cdots \pi_n \in \mathfrak{S}_n$ such that the map $w : P \to [n]$ defined by $w(\omega^{-1}(\pi_i)) = i$ is a linear extension of P. Thus $\mathcal{L}(P, \omega)$ may be regarded as the set of linear extensions of P, regarded as permutations of the labels of P. For instance, if (P, ω) is given by



then $\mathcal{L}(P, \omega) = \{52143, 52413, 25143, 25413, 24513, 52134, 25134\}$.

The following result is the "fundamental theorem of (reverse) (P, ω)-partitions." The proof is exactly like that of the special case Lemma 4.5.3 (in its "reverse form") and will be omitted.

7.19.4 Theorem. *Let $\mathcal{A}^r(P, \omega)$ denote the set of all reverse (P, ω)-partitions. Then*

$$\mathcal{A}^r(P, \omega) = \bigcup_{\pi \in \mathcal{L}(P,\omega)}^{\cdot} S_\pi^r \qquad \text{(disjoint union).}$$

In exact analogy to the definition (7.92) of $K_P(x)$, let

$$K_{P,\omega}(x) = \sum_{\sigma \in \mathcal{A}^r(P,\omega)} x^\sigma .$$

Note that $K_{P,\omega}(x)$ is quasisymmetric. Just as (7.94) follows from (7.93), we obtain:

7.19.5 Corollary. *We have*

$$K_{P,\omega} = \sum_{\pi \in \mathcal{L}(P,\omega)} L_{\mathrm{co}(\pi)}, \tag{7.95}$$

the expansion of $K_{P,\omega}$ in terms of the fundamental quasisymmetric functions.

We now want to apply Corollary 7.19.5 to the skew Schur functions $s_{\lambda/\mu}$, where $|\lambda/\mu| = n$. Define $P_{\lambda/\mu}$ to be the poset whose elements are the squares (i, j) of λ/μ, partially ordered componentwise. Thus $P_{\lambda/\mu}$ is a finite convex subset of $\mathbb{P} \times \mathbb{P}$, and every finite convex subset of $\mathbb{P} \times \mathbb{P}$ is equal to $P_{\lambda/\mu}$ for some λ/μ. (The case $\mu = \emptyset$ was discussed already in Proposition 7.10.3(b).) Define a labeling $\omega_{\lambda/\mu} : P_{\lambda/\mu} \to [n]$, called the *Schur labeling* of $P_{\lambda/\mu}$, as follows: The bottom square of the first column of $P_{\lambda/\mu}$ is labeled 1. The labeling then proceeds up the first column, then up the second column, etc. For instance, Figure 7-8 shows the Schur labeling of the skew shape 5331/2, drawn both as a Young diagram and as a labeled poset.

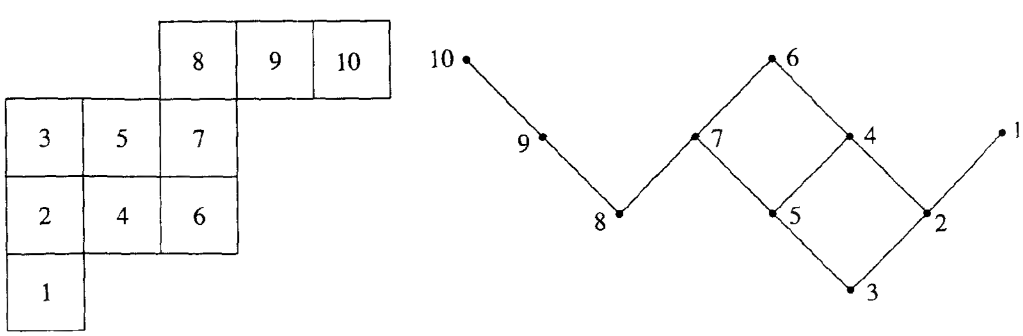


Figure 7-8. A Young diagram and the corresponding Schur labeled poset.

It is immediate from the definition of $K_{P,\omega}$ that a reverse $(P_{\lambda/\mu}, \omega_{\lambda/\mu})$-partition is just an SSYT, so

$$K_{P_{\lambda/\mu}, \omega_{\lambda/\mu}} = s_{\lambda/\mu}.$$

Hence as a special case of Corollary 7.19.5 we obtain the expansion of $s_{\lambda/\mu}$ in terms of fundamental quasisymmetric functions. Rather than describe this expansion directly in terms of $\mathcal{L}(P_{\lambda/\mu}, \omega_{\lambda/\mu})$ as in Corollary 7.19.5, we want a description in terms of SYTs of shape λ/μ.

A linear extension $w : P_{\lambda/\mu} \to [n]$ corresponds to an SYT T_w of shape λ/μ. Similarly w corresponds to a permutation $\pi_w \in \mathcal{L}(P_{\lambda/\mu}, \omega_{\lambda/\mu})$. Define a *descent* of an SYT T to be an integer i such that $i + 1$ appears in a lower row of T than i, and define the *descent set* $D(T)$ to be the set of all descents of T. For instance, the SYT

$$\begin{array}{cccccc} & & & 2 & 8 \\ & & 1 & 4 & 5 & 10 \\ 3 & 6 & 9 \\ 7 \end{array}$$

has descent set $\{2, 5, 6, 8\}$. We write $\mathrm{co}(T)$ for the composition $\mathrm{co}(D(T))$ of n associated with the descent set $D(T)$.

7.19.6 Lemma. *Let $w : P_{\lambda/\mu} \to [n]$ be a linear extension. Then $D(T_w) = D(\pi_w)$.*

Proof. Let $1 \leq i \leq n - 1$. Let $s = (a, b)$ be the square of T_w containing i. The square $s' = (a', b')$ containing $i + 1$ satisfies either (a) $a' \leq a$ and $b' > b$, or (b) $a' > a$ and $b' \leq b$. In the former case, $i \notin D(T_w)$ and $\omega(s') > \omega(s)$, where $\omega = \omega_{\lambda/\mu}$. Hence also $i \notin D(\pi_w)$. In the latter case, $i \in D(T_w)$ and $\omega(s') < \omega(s)$. Hence also $i \in D(\pi_w)$, and the proof follows. □

Combining Corollary 7.19.5 and Lemma 7.19.6 gives the main result of this section.

7.19.7 Theorem. *We have*

$$s_{\lambda/\mu} = \sum_T L_{\mathrm{co}(T)},$$

where T ranges over all SYTs of shape λ/μ.

7.19.8 Example. Let $\lambda/\mu = 32/1$. The five SYTs of shape λ/μ, with their descents shown in boldface, are:

$$\begin{array}{ccccc} \begin{array}{cc} \mathbf{1} & 4 \\ 2 & 3 \end{array} & \begin{array}{cc} 1 & \mathbf{2} \\ 3 & 4 \end{array} & \begin{array}{cc} \mathbf{2} & 4 \\ 1 & 3 \end{array} & \begin{array}{cc} 2 & \mathbf{3} \\ 1 & 4 \end{array} & \begin{array}{cc} 1 & \mathbf{3} \\ 2 & 4 \end{array} \end{array}$$

Hence

$$s_{32/1} = L_{13} + 2L_{22} + L_{31} + L_{121}.$$

As an illustration of the use of Theorem 7.19.7, we have the following somewhat surprising combinatorial result. See Exercise 7.90(b) for a more direct proof.

7.19.9 Proposition. *Let $|\lambda/\mu| = n$. For any $1 \le i \le n-1$, the number $d_i(\lambda/\mu)$ of SYTs T of shape λ/μ for which $i \in D(T)$ is independent of i.*

Proof. Define a linear transformation $\varphi_i : \mathcal{Q}^n \to \mathbb{Q}$ by

$$\varphi_i(L_\alpha) = \begin{cases} 1 & i \in S_\alpha \\ 0 & \text{otherwise.} \end{cases}$$

By Theorem 7.19.7 we have $d_i(\lambda/\mu) = \varphi_i(s_{\lambda/\mu})$. Hence it suffices to prove that $\varphi_i(b_\nu)$ is independent of i for some basis $\{b_\nu\}$ of Λ^n, for then $\varphi_i(f)$ will be independent of i for all $f \in \Lambda^n$, including $f = s_{\lambda/\mu}$. We choose $b_\nu = m_\nu$.

By definition of M_α we have

$$m_\nu = \sum_\alpha M_\alpha,$$

where α ranges over all distinct permutations of the parts of ν. If $S \subseteq [n-1]$ and $\#S \le n-3$, then every $i \in [n-1]$ appears in the same number of sets T of even cardinality satisfying $S \subseteq T \subseteq [n-1]$ as of odd cardinality. It follows from (7.91) that $\varphi_i(M_\alpha) = 0$ if $\ell(\alpha) \le n-2$ (i.e., $\#S_\alpha \le n-3$), so $\varphi_i(m_\nu) = 0$ (independent of i) unless possibly $\nu = \langle 1^n \rangle$ or $\nu = \langle 21^{n-2} \rangle$.

If $\nu = \langle 1^n \rangle$, then $m_{1^n} = L_{1,1,\dots,1}$, so $\varphi_i(m_{1^n}) = 1$ (independent of i). If $\nu = \langle 21^{n-2} \rangle$, then by (7.91) we have

$$m_{21^{n-2}} = \sum_{j=1}^{n-1} (L_{1^{j-1},2,1^{n-j-1}} - L_{1,1,\dots,1}).$$

It follows that $\varphi_i(m_{21^{n-2}}) = -1$ (independent of i), and the proof follows. □

If we write $s_{\lambda/\mu} = \sum_\nu K_{\lambda/\mu,\nu} m_\nu$ and apply φ_i, then the above proof shows that

$$\begin{aligned} \varphi_i(s_{\lambda/\mu}) &= K_{\lambda/\mu,1^n} - K_{\lambda/\mu,21^{n-2}} \\ &= f^{\lambda/\mu} - K_{\lambda/\mu,21^{n-2}}. \end{aligned}$$

It is easy to see that when $\mu = \emptyset$ this quantity is equal to $f^{\lambda/11}$. Alternatively, it is clear that $\varphi_1(s_\lambda) = f^{\lambda/11}$, since if T is an SYT of shape λ and $1 \in D(T)$, then T has a 1 in the $(1,1)$ square and a 2 in the $(2,1)$ square. Hence, given that $\varphi_i(s_\lambda)$ is independent of i, there follows $\varphi_i(s_\lambda) = f^{\lambda/11}$ as before.

As another and more significant application of Theorem 7.19.7, we give a formula for the stable principal specialization $s_{\lambda/\mu}(1, q, q^2, \ldots)$.

7.19.10 Lemma. *Let* $\alpha \in \text{Comp}(n)$. *Then*

$$L_\alpha(1, q, q^2, \ldots) = \frac{q^{e(\alpha)}}{(1-q)(1-q^2)\cdots(1-q^n)},$$

where $e(\alpha) = \sum_{i \in S_\alpha}(n-i)$.

Proof. By (7.89) we have

$$L_\alpha(1, q, q^2, \ldots) = \sum_{\substack{i_1 \le \cdots \le i_n \\ i_j < i_{j+1} \text{ if } j \in S_\alpha}} q^{i_1+i_2+\cdots+i_n-n}.$$

Define

$$r_j = i_j - 1 - \#\{m \in S_\alpha : m < j\}.$$

Then

$$L_\alpha(1, q, q^2, \ldots) = q^{i(\alpha)} \sum_{0 \le r_1 \le \cdots \le r_n} q^{r_1+r_2+\cdots+r_n},$$

where

$$i(\alpha) = \sum_{j=1}^{n} \#\{m \in S_\alpha : m < j\}.$$

It is easy to see that $i(\alpha) = e(\alpha)$, and the proof follows. □

For any SYT T define the *major index* * $\text{maj}(T)$ by

$$\text{maj}(T) = \sum_{i \in D(T)} i.$$

7.19.11 Proposition. *Let* $|\lambda/\mu| = n$. *Then*

$$s_{\lambda/\mu}(1, q, q^2, \ldots) = \frac{\sum_T q^{\text{maj}(T)}}{(1-q)(1-q^2)\cdots(1-q^n)},$$

where T ranges over all SYTs of shape λ/μ.

Proof. Combining Theorem 7.19.7 and Lemma 7.19.10 yields

$$s_{\lambda/\mu}(1, q, q^2, \ldots) = \frac{\sum_T q^{\text{comaj}(T)}}{(1-q)(1-q^2)\cdots(1-q^n)},$$

* In Section 4.5 we used the notation $\iota(\pi)$ rather than $\text{maj}(\pi)$ for the analogous concept of the major index (or greater index) of a permutation π.

where $\mathrm{comaj}(T) = e(D(T)) = \sum_{i \in D(T)} (n-i)$, the *comajor index* of T. Let $s_{\lambda/\mu} = \sum_\alpha c_\alpha L_\alpha$. By Proposition 7.19.2 we see that $c_\alpha = c_{\alpha^*}$, where if $\alpha = (\alpha_1, \ldots, \alpha_k)$ then $\alpha^* = (\alpha_k, \ldots, \alpha_1)$. Hence

$$\sum_T q^{\mathrm{comaj}(T)} = \sum_T q^{\mathrm{maj}(T)},$$

summed over all SYT of shape λ/μ, and the proof follows. □

Suppose that λ/μ is a disjoint union of squares, e.g., 4321/321 is a disjoint union of four squares. Then the SYTs T of shape λ/μ correspond in a natural way to permutations $\pi \in \mathfrak{S}_n$ such that $D(T) = D(\pi)$. Clearly $s_{\lambda/\mu}(1, q, q^2, \ldots) = (1-q)^{-n}$, so Proposition 7.19.11 reduces to Corollary 4.5.9, viz.,

$$\sum_{\pi \in \mathfrak{S}_n} q^{\mathrm{maj}(\pi)} = (1+q)(1+q+q^2)\cdots(1+q+\cdots+q^{n-1}).$$

In Corollary 7.21.3 we will evaluate $s_\lambda(1, q, q^2, \ldots)$ explicitly, thereby yielding an explicit formula (Corollary 7.21.5) for $\sum_T q^{\mathrm{maj}(T)}$, summed over all SYTs of shape λ. Another formula for $s_{\lambda/\mu}(1, q, q^2, \ldots)$, similar to Proposition 7.19.11 but with a different denominator, is given by Exercise 7.102(b).

As a variant of Proposition 7.19.11, we can find the *number* of descents of T, rather than the sum of the descents. Let $d(T) = \#D(T)$. For a power series $f(x_1, x_2, \ldots)$ write $f(1^m) = f(x_1 = \cdots = x_m = 1,\ x_{m+1} = x_{m+2} = \cdots = 0)$, as in equation (7.8). For $\alpha \in \mathrm{Comp}(n)$ it is easy to see that

$$L_\alpha(1^m) = \left(\!\!\binom{m - \#S_\alpha}{n}\!\!\right) = \binom{m - \#S_\alpha + n - 1}{n}.$$

It now follows immediately from Theorem 7.19.7 (analogously to Theorem 4.5.14) that if $|\lambda/\mu| = n$ then

$$\sum_{m \geq 0} s_{\lambda/\mu}(1^m) t^m = \frac{\sum_T t^{d(T)+1}}{(1-t)^{n+1}}. \tag{7.96}$$

For instance, if λ is the "hook" $k1^{n-k}$, then all $\binom{n-1}{k-1}$ SYTs of shape λ have $n-k$ descents, from which there follows

$$s_{k1^{n-k}}(1^m) = \binom{n-1}{k-1}\binom{m+k-1}{n}.$$

Similarly we can take into account *both* $d(T)$ and $\mathrm{maj}(T)$. We merely state the resulting formula, whose proof is analogous to that of Exercise 4.24(b).

7.19.12 Proposition. *We have*

$$s_{\lambda/\mu}(1, q, \ldots, q^{m-1}) = \sum_T \binom{m - d(T) + n - 1}{n} q^{\mathrm{maj}(T)},$$

where T ranges over all SYTs of shape λ/μ.

An explicit formula for $s_\lambda(1, q, \ldots, q^{m-1})$ is given by Theorem 7.21.2.

7.20 Plane Partitions and the RSK Algorithm

We have now developed enough of the theory of symmetric functions that we can give a number of enumerative applications. This section and the next two will be devoted to a fascinating generalization of partitions of integers known as "plane partitions." A *plane partition* is an array $\pi = (\pi_{ij})_{i,j\geq 1}$ of nonnegative integers such that π has finite support (i.e., finitely many nonzero entries) and is weakly decreasing in rows and columns. If $\sum \pi_{ij} = n$, then we write $|\pi| = n$ and say that π is a plane partition *of* n. When writing examples of plane partitions, the 0's (or all but finitely many 0's) are suppressed. Thus the plane partitions of integers $0 \leq n \leq 3$ are given by

$$\begin{array}{ccccccccccc} \emptyset & 1 & 2 & 11 & 1 & 3 & 21 & 111 & 11 & 2 & 1 \\ & & & & 1 & & & & 1 & 1 & 1 \\ & & & & & & & & & & 1. \end{array}$$

An ordinary partition $\lambda \vdash n$ may be regarded as a weakly decreasing *one-dimensional* array $(\lambda_1, \lambda_2, \ldots)$ of nonnegative integers with finite support. Thus plane partitions are a natural generalization to two dimensions of ordinary partitions. It now seems obvious to define *r-dimensional partitions* for any $r \geq 1$. However, almost nothing significant is known for $r \geq 3$. Plane partitions have obvious similarities with semistandard tableaux. Indeed, a reverse SSYT is just a special kind of plane partition (with the irrelevant 0's removed), and in fact in our definition of reverse SSYT we mentioned the alternative term "column-strict plane partition." Because of the similarity between SSYTs and plane partitions, it is not surprising that symmetric functions play an important role in the enumeration of plane partitions.

A *part* of a plane partition $\pi = (\pi_{ij})$ is a positive entry $\pi_{ij} > 0$. The *shape* of π is the ordinary partition λ for which π has λ_i nonzero parts in the i-th row (so $\pi_{i\lambda_i} > 0$, $\pi_{i,\lambda_i+1} = 0$). We say that π has r *rows* if $r = \ell(\lambda)$. Similarly, π has s *columns* if $s = \ell(\lambda') = \lambda_1$. Write $\ell_1(\pi)$ for the number of rows and $\ell_2(\pi)$ for the number of columns of π. Finally, define the *trace* of π by the usual formula $\mathrm{tr}(\pi) = \sum \pi_{ii}$. For example, the plane partition

$$\begin{array}{cccccccc} 7 & 5 & 5 & 3 & 2 & 1 & 1 & 1 \\ 6 & 5 & 5 & 2 & 1 & 1 \\ 6 & 3 & 2 & 2 \end{array}$$

has shape $(8, 6, 4)$, 18 parts, 3 rows, 8 columns, and trace 14.

Let $\mathcal{P}(r, c)$ be the set of all plane partitions with at most r rows and at most c columns. For instance, if $\pi \in \mathcal{P}(1, c)$, then π is just an ordinary partition and $\mathrm{tr}(\pi)$ is the largest part of π. It is then clear by "inspection" (looking at the conjugate

partition π' instead of π) that

$$\sum_{\pi \in \mathcal{P}(1,c)} q^{\operatorname{tr}(\pi)} x^{|\pi|} = \frac{1}{(1-qx)(1-qx^2)\cdots(1-qx^c)}. \tag{7.97}$$

The main result of this section is the following generalization of equation (7.97).

7.20.1 Theorem. *Fix $r, c \in \mathbb{P}$. Then*

$$\sum_{\pi \in \mathcal{P}(r,c)} q^{\operatorname{tr}(\pi)} x^{|\pi|} = \prod_{i=1}^{r} \prod_{j=1}^{c} (1 - qx^{i+j-1})^{-1}.$$

Proof. We will give an elegant bijective proof based on the RSK algorithm and a simple method of merging a pair of reverse SSYT of the same shape into a single plane partition. First we describe how to merge two partitions λ and μ with distinct parts and with the same number of parts into a single partition $\rho = \rho(\lambda, \mu)$. Draw the Ferrers diagram of λ but with each row indented one space to the right of the beginning of the previous row. Such a diagram is called the *shifted* Ferrers diagram of λ. For instance, if $\lambda = (5, 3, 2)$ then we get the shifted diagram

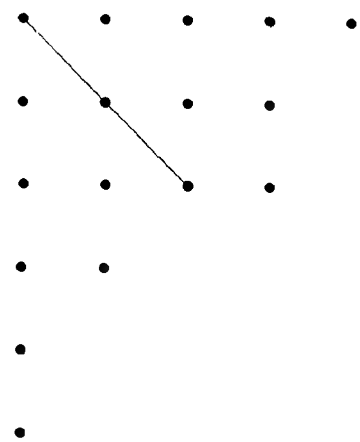
$$\begin{array}{ccccc}
\bullet & \bullet & \bullet & \bullet & \bullet \\
 & \bullet & \bullet & \bullet & \\
 & & \bullet & \bullet &
\end{array}$$

Do the same for μ, and then transpose the diagram. For instance, if $\mu = (6, 3, 1)$ then we get the transposed shifted diagram

$$\begin{array}{ccc}
\bullet & & \\
\bullet & \bullet & \\
\bullet & \bullet & \bullet \\
\bullet & \bullet & \\
\bullet & & \\
\bullet & &
\end{array}$$

Now merge the two diagrams into a single diagram by identifying their main diagonals. For λ and μ as above, we get the diagram (with the main diagonal

drawn for clarity)



Define $\rho(\lambda, \mu)$ to be the partition for which this merged diagram is the Ferrers diagram. The above example shows that

$$\rho(532, 631) = 544211.$$

The map $(\lambda, \mu) \mapsto \rho(\lambda, \mu)$ is clearly a bijection between pairs of partitions (λ, μ) with k distinct parts, and partitions ρ of rank k (as defined in Section 7.2). Note that

$$|\rho| = |\lambda| + |\mu| - \ell(\lambda).$$

We now extend the above bijection to pairs (P, Q) of reverse SSYTs of the same shape. If λ^i denotes the i-th column of P and μ^i the i-th column of Q, then let $\pi(P, Q)$ be the array whose i-th column is $\rho(\lambda^i, \mu^i)$. For instance, if

$$P = \begin{array}{llll} 4 & 4 & 2 & 1 \\ 3 & 1 & 1 & \\ 2 & & & \end{array} \quad \text{and} \quad Q = \begin{array}{llll} 5 & 3 & 2 & 2 \\ 4 & 2 & 1 & \\ 1 & & & \end{array} ,$$

then

$$\pi(P, Q) = \begin{array}{llll} 4 & 4 & 2 & 1 \\ 4 & 2 & 2 & 1 \\ 4 & 2 & & \\ 2 & & & \\ 2 & & & \end{array} .$$

It is easy to see that $\pi(P, Q)$ is a plane partition. Replace each row of $\pi(P, Q)$ by its conjugate to obtain another plane partition $\pi'(P, Q)$. With $\pi(P, Q)$ as above

we obtain

$$\begin{array}{rl} & 4\,3\,2\,2 \\ & 4\,3\,1\,1 \\ \pi'(P, Q) = & 2\,2\,1\,1\,. \\ & 1\,1 \\ & 1\,1 \end{array}$$

One can easily check that the map $(P, Q) \mapsto \pi'(P, Q)$ is a bijection between pairs (P, Q) of reverse SSYTs of the same shape and plane partitions π'. Write $\operatorname{diag}(\pi')$ for the main diagonal $(\pi'_{11}, \pi'_{22}, \dots)$ of π', $\max(P)$ for the largest part P_{11} of the reverse SSYT P, etc. Recall that $\operatorname{sh}(P)$ denotes the shape of P, so $\operatorname{sh}(P) = \operatorname{sh}(Q)$, with P, Q as above. It is easy to see that

$$\begin{aligned} |\pi'| &= |P| + |Q| - |\operatorname{sh}(P)| \\ \operatorname{diag}(\pi') &= \operatorname{sh}(P) = \operatorname{sh}(Q), \qquad \text{so } \operatorname{tr}(\pi') = |\operatorname{sh}(P)| \\ \ell_1(\pi') &= \max(Q) \\ \ell_2(\pi') &= \max(P). \end{aligned} \tag{7.98}$$

Now let $A = (a_{ij})$ be an $\mathbb{N}$-matrix of finite support. We want to associate with A a pair of *reverse* SSYTs of the same shape. This can be done by an obvious variant of the RSK algorithm, where we reverse the roles of $\leq$ and $\geq$ in defining row insertion. Equivalently, if

$$w_A = \begin{pmatrix} u_1 & \cdots & u_n \\ v_1 & \cdots & v_n \end{pmatrix}$$

is the two-line array associated with A, then apply the ordinary RSK algorithm to the two-line array,

$$\begin{pmatrix} -u_n & \cdots & -u_1 \\ -v_n & \cdots & -v_1 \end{pmatrix}$$

(whose entries are now *negative* integers) and then change the sign back to positive of all entries of the pair of SSYTs. We will obtain a pair (P, Q) of reverse SSYTs satisfying

$$\begin{aligned} |P| &= \sum_{i,j} j a_{ij} \\ |Q| &= \sum_{i,j} i a_{ij} \\ \max(P) &= \max\{j : a_{ij} \neq 0\} \\ \max(Q) &= \max\{i : a_{ij} \neq 0\} \\ |\operatorname{sh}(P)| &= |\operatorname{sh}(Q)| = \sum a_{ij}. \end{aligned} \tag{7.99}$$

It follows from the equations beginning with (7.98) and (7.99) that if $\mathcal{M}_{rc}$ is the set of all $r \times c$ $\mathbb{N}$-matrices, then

$$\sum_{\pi \in \mathcal{P}(r,c)} q^{\operatorname{tr}(\pi)} x^{|\pi|} = \sum_{A=(a_{ij}) \in \mathcal{M}_{rc}} q^{\sum a_{ij}} x^{\sum (i+j) a_{ij} - \sum a_{ij}}$$

$$= \prod_{i=1}^{r} \prod_{j=1}^{c} \left(\sum_{a_{ij} \geq 0} q^{a_{ij}} x^{(i+j-1) a_{ij}} \right)$$

$$= \prod_{i=1}^{r} \prod_{j=1}^{c} (1 - q x^{i+j-1})^{-1}. \qquad \square$$

Now let $\mathcal{P}(r)$ denote the set of all plane partitions with at most r rows. If we let $q = 1$ and $c \to \infty$ in Theorem 7.20.1, then we obtain the following elegant enumeration of the elements of $\mathcal{P}(r)$.

7.20.2 Corollary. *Fix $r \in \mathbb{P}$. Then*

$$\sum_{\pi \in \mathcal{P}(r)} x^{|\pi|} = \prod_{i \geq 1} (1 - x^i)^{-\min(i,r)}. \tag{7.100}$$

Proof. Theorem 7.20.1 yields

$$\sum_{\pi \in \mathcal{P}(r)} x^{|\pi|} = \prod_{i=1}^{r} \prod_{j \geq 1} (1 - x^{i+j-1})^{-1},$$

which is easily seen to agree with the right-hand side of (7.100). $\square$

Finally, let $\mathcal{P}$ denote the set of all plane partitions, and let $r \to \infty$ in Corollary 7.20.2 to obtain the archetypal result in the theory of plane partitions:

7.20.3 Corollary. *We have*

$$\sum_{\pi \in \mathcal{P}} x^{|\pi|} = \prod_{i \geq 1} (1 - x^i)^{-i}.$$

A nice variation of Theorem 7.20.1 arises when we take into account the symmetry result Theorem 7.13.1 of the RSK algorithm. Define a plane partition $\sigma = (\sigma_{ij})$ to be *symmetric* if $\sigma_{ij} = \sigma_{ji}$ for all i, j. Let $\mathcal{S}(r)$ denote the set of all symmetric plane partitions with at most r rows (and therefore with at most r columns).

7.20.4 Theorem. *Fix $r \in \mathbb{P}$. Then*

$$\sum_{\sigma \in \mathcal{S}(r)} q^{\mathrm{tr}(\sigma)} x^{|\sigma|} = \prod_{i=1}^{r} (1 - qx^{2i-1})^{-1} \cdot \prod_{1 \le i < j \le r} \left(1 - q^2 x^{2(i+j-1)}\right)^{-1}.$$

First Proof. Let $\pi'(P, Q)$ be the plane partition described in the proof of Theorem 7.20.1. It is easy to see that π' is symmetric if and only if $P = Q$. Moreover, suppose that $A \xrightarrow{\mathrm{RSK}'} (P, Q)$, where RSK′ is the "reverse" RSK algorithm also described in the proof of Theorem 7.20.1 (so P and Q are reverse SSYTs). Theorem 7.13.1 holds equally well for RSK′, so $A \xrightarrow{\mathrm{RSK}'} (P, P)$ if and only if $A = A'$. Let $\mathcal{M}'_r$ be the set of all $r \times r$ symmetric $\mathbb{N}$-matrices. We then obtain, exactly as in the proof of Theorem 7.20.1, that

$$\begin{aligned}
\sum_{\sigma \in \mathcal{S}(r)} q^{\mathrm{tr}(\sigma)} x^{|\sigma|} &= \sum_{A=(a_{ij}) \in \mathcal{M}'_r} q^{\sum a_{ij}} x^{\sum (i+j-1) a_{ij}} \\
&= \sum_{A} q^{\sum_{i \ge 1} a_{ii} + 2 \sum_{1 \le i < j \le r} a_{ij}} x^{\sum_{i \ge 1} (2i-1) a_{ii} + 2 \sum_{1 \le i < j \le r} (i+j-1) a_{ij}} \\
&= \prod_{i=1}^{r} \left(\sum_{a_{ii} \ge 0} q^{a_{ii}} x^{(2i-1) a_{ii}} \right) \cdot \prod_{1 \le i < j \le r} \left(\sum_{a_{ij} \ge 0} q^{2a_{ij}} x^{2(i+j-1) a_{ij}} \right) \\
&= \prod_{i=1}^{r} (1 - qx^{2i-1})^{-1} \cdot \prod_{1 \le i < j \le r} \left(1 - q^2 x^{2(i+j-1)}\right)^{-1}. \qquad \square
\end{aligned}$$

Second Proof. There is a clever alternative proof that avoids the use of Theorem 7.13.1. Suppose that π is a reverse SSYT with only odd parts. Thus each column of π is a partition into distinct odd parts. The solution to Exercise 1.22(d) gives a bijection between partitions of n with distinct odd parts and self-conjugate partitions of n. Apply this bijection to each column of π, and then take the conjugate of each row, producing a symmetric plane partition σ. An example is given by

$$\pi = \begin{matrix} 7\,5\,5\,5\,5\,1\,1 \\ 5\,3\,3\,1\,1 \\ 1\,1 \end{matrix} \quad \to \quad \begin{matrix} 4\,3\,3\,3\,3\,1\,1 \\ 4\,3\,3\,2\,2 \\ 3\,3\,2\,1\,1 \\ 2 \end{matrix} \quad \to \quad \begin{matrix} 7\,5\,5\,1 \\ 5\,5\,3\,1 \\ 5\,3\,2 \\ 1\,1 \end{matrix} = \sigma.$$

One can easily check that $|\pi| = |\sigma|$, $\mathrm{sh}(\pi) = \mathrm{diag}(\sigma)$, $|\mathrm{sh}(\pi)| = \mathrm{tr}(\sigma)$, and $\ell_1(\sigma) = \frac{1}{2}[1 + \max(\pi)]$. Hence from Corollary 7.13.8 there follows

$$\begin{aligned}
\sum_{\sigma \in \mathcal{S}(r)} q^{\mathrm{tr}(\sigma)} x^{|\sigma|} &= \prod_{\substack{i=1 \\ i \text{ odd}}}^{2r-1} (1 - x_i)^{-1} \cdot \prod_{\substack{1 \le i < j \le 2r-1 \\ i,j \text{ odd}}} (1 - x_i x_j)^{-1} \Big|_{x_i = qx^i} \\
&= \prod_{i=1}^{r} (1 - qx^{2i-1})^{-1} \cdot \prod_{1 \le i < j \le r} \left(1 - q^2 x^{2(i+j-1)}\right)^{-1}. \qquad \square
\end{aligned}$$

Putting $q = 1$ and letting $r \to \infty$ yields the following elegant enumeration of all symmetric plane partitions of n.

7.20.5 Corollary. *Let $\mathcal{S}$ be the set of all symmetric plane partitions. Then*

$$\sum_{\sigma \in \mathcal{S}} x^{|\sigma|} = \prod_{i \geq 1} \frac{1}{(1 - x^{2i-1})(1 - x^{2i})^{\lfloor i/2 \rfloor}}.$$

7.21 Plane Partitions with Bounded Part Size

Our main object in this section is to refine Theorem 7.20.1, in the case $q = 1$, by restricting the size of the largest part of the plane partition $\pi \in \mathcal{P}(r, c)$. Consider, for instance, the special case $r = 1$, so that π is just an ordinary partition $\lambda = (\lambda_1, \ldots, \lambda_c)$ with at most c parts. If we add the restriction $\lambda_1 \leq t$, then Proposition 1.3.19 tells us that

$$\sum_{\substack{\lambda = (\lambda_1, \ldots, \lambda_c) \\ \lambda_1 \leq t}} q^{|\lambda|} = \binom{c+t}{c},$$

a q-binomial coefficient. It is this result that we wish to generalize to plane partitions. We cannot expect a nice bijective proof like that of Theorem 7.20.1, because even in the case $r = 1$ the expansion of the numerator $(1 - q^{c+t}) \cdots (1 - q^{t+1})$ of the q-binomial coefficient $\binom{c+t}{c}$ has negative coefficients. A bijective proof would have to involve either an involution principle argument (or something similar), or else moving the numerator of $\binom{c+t}{c}$ over to the other side. While such proofs do exist, they lack the elegance of the proof of Theorem 7.20.1. The proof we give here will not be bijective, but will be a simple consequence of symmetric function theory.

To understand better the significance of the restrictions on the number of rows, the number of columns, and the largest part, we first discuss the notion of the diagram of a plane partition, generalizing the notion of the Young or Ferrers diagram of a partition. Formally, the *diagram* $D(\pi)$ (often identified with π) of a plane partition $\pi = (\pi_{ij})$ is the subset of $\mathbb{P}^3$ defined by

$$D(\pi) = \{(i, j, k) \in \mathbb{P}^3 : 1 \leq k \leq \pi_{ij}\}.$$

Think of replacing the entry π_{ij} by a pillar of π_{ij} cubes (or dots). For instance, the (Ferrers) diagram of the plane partition

$$\pi = \begin{matrix} 4 & 2 & 1 \\ 3 & 1 & \\ 1 & & \\ 1 & & \end{matrix}$$

is given by Figure 7-9.

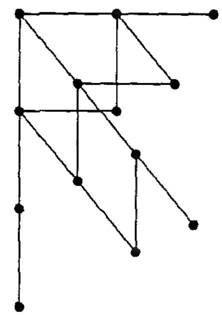


Figure 7-9. The diagram of a plane partition.

Any permutation w of the three coordinate axes transforms (the diagram of) a plane partition π of n into another plane partition $w(\pi)$ of n. Thus a plane partition has *six* "associates," called *aspects*, indexed by elements of $\mathfrak{S}_3$. Compare with the two "associates" λ and λ' of an ordinary partition λ. In terms of the plane partition $\pi = (\pi_{ij})$ itself, the six aspects are obtained as follows:

- leave π unchanged,
- conjugate every row of π,
- conjugate every column of π,
- transpose π,
- conjugate every row of π and then transpose,
- conjugate every column of π and then transpose.

The three statistics $\ell_1(\pi)$, $\ell_2(\pi)$, and $\max(\pi)$ are permuted among themselves when we take an aspect $w(\pi)$ of π. Thus for instance the number of plane partitions of n with at most r rows and at most c columns equals the number of plane partitions of n with at most c rows and with largest part at most r. Since we have enumerated plane partitions of n with at most r rows and at most c columns (Theorem 7.20.1 when $q = 1$), it now seems very natural to consider an additional restriction on the largest part.

Let $r, c, t \in \mathbb{P}$, and define the *box*

$$B(r, c, t) = \{(i, j, k) \in \mathbb{P}^3 : 1 \le i \le r,\ 1 \le j \le c,\ 1 \le k \le t\}.$$

Thus a plane partition π satisfies $\ell_1(\pi) \le r$, $\ell_2(\pi) \le c$, and $\max(\pi) \le t$ if and only if its diagram is contained in the box $B(r, c, t)$, which we write as $\pi \subseteq B(r, c, t)$. Our current goal, then, is to evaluate the generating function $\sum_{\pi \subseteq B(r,c,t)} q^{|\pi|}$.

As a preliminary step we will evaluate the principal specialization $s_\lambda(1, q, \dots, q^{n-1})$. The most elegant formulation of this result involves two important statistics associated with the boxes of the Young diagram of a partition. Given a Young

diagram λ (where we are identifying the diagram $\{(i,j) : 1 \le j \le \lambda_i\}$ with its shape) and a square $u = (i,j) \in \lambda$, define the *hook length* $h(u)$ *of* λ *at* u by

$$h(u) = \lambda_i + \lambda'_j - i - j + 1.$$

Equivalently, $h(u)$ is the number of squares directly to the right or directly below u, counting u itself once. For instance, the partition 4421 has hook lengths given by

$$\begin{array}{|c|c|c|c|} \hline 7 & 5 & 3 & 2 \\ \hline 6 & 4 & 2 & 1 \\ \hline 3 & 1 \\ \cline{1-2} 1 \\ \cline{1-1} \end{array}$$

Similarly define the *content* $c(u)$ of λ at $u = (i,j)$ by

$$c(u) = j - i.$$

For $\lambda = 4421$ the contents are given by

$$\begin{array}{|c|c|c|c|} \hline 0 & 1 & 2 & 3 \\ \hline -1 & 0 & 1 & 2 \\ \hline -2 & -1 \\ \cline{1-2} -3 \\ \cline{1-1} \end{array}$$

7.21.1 Lemma. *Let* $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathrm{Par}$ *and* $\mu_i = \lambda_i + n - i$. *Then*

$$\prod_{u \in \lambda} [h(u)] = \frac{\prod_{i \ge 1} [\mu_i]!}{\prod_{1 \le i < j \le n} [\mu_i - \mu_j]} \tag{7.101}$$

$$\prod_{u \in \lambda} [n + c(u)] = \prod_{i=1}^{n} \frac{[\mu_i]!}{[n-i]!}, \tag{7.102}$$

where $[k] = 1 - q^k$ *and* $[k]! = [1][2]\cdots[k]$.

Proof. Trusting that "one picture is worth a thousand words," we will illustrate the proofs with an example. Let $\lambda = 4421$. Add $n - i$ squares to the i-th row of the diagram of λ, obtaining the diagram of μ. In square (i,j) insert the number $\mu_i - j + 1$. Thus the multiset of inserted numbers is just $\bigcup_{i \ge 1} \{1, 2, \ldots, \mu_i\}$, the exponents in the numerator of the right-hand side of equation (7.101) (when written as a product of factors $1 - q^k$). For each $1 \le i < j \le n$, write the number

$\mu_i - \mu_j$ in square $(i, \mu_j + 1)$ in boldface. We obtain the array

7	**6**	5	**4**	3	2	**1**
6	**5**	4	**3**	2	1	
3	**2**	1				
1						

A little thought shows that if we remove the columns $\mu_j + 1$ of bold numbers, we obtain just the diagram of λ with the hook length $h(u)$ in square u. This proves (7.101).

An analogous (but even simpler) argument works for equation (7.102). Here the relevant array is

1	**2**	**3**	4	5	6	7
	1	**2**	3	4	5	6
		1	2	3		
			1			

This completes the proof. $\square$

Given $\lambda \vdash n$, define

$$b(\lambda) = \sum (i-1)\lambda_i = \sum \binom{\lambda_i'}{2}. \tag{7.103}$$

Note that $b(\lambda)$ is the smallest possible sum of the entries of an SSYT (allowing 0 as a part) of shape λ, obtained uniquely by placing $i - 1$ in all the squares of the i-th row of λ. In particular, for $n \geq \ell(\lambda)$ we have $s_\lambda(1, q, \dots, q^{n-1}) = q^{b(\lambda)} v_\lambda(q)$, where $v_\lambda(q)$ is a polynomial in q satisfying $v_\lambda(0) = 1$. (If $n < \ell(\lambda)$ then $s_\lambda(1, q, \dots, q^{n-1}) = 0$.)

7.21.2 Theorem. *For any $\lambda \in$ Par and $n \in \mathbb{P}$ we have*

$$s_\lambda(1, q, \dots, q^{n-1}) = q^{b(\lambda)} \prod_{u \in \lambda} \frac{[n + c(u)]}{[h(u)]}.$$

Proof. If $n < \ell(\lambda)$ then both sides vanish, so assume $n \geq \ell(\lambda)$. By Theorem 7.15.1 (the bialternant formula for $s_\lambda(x_1, \dots, x_n)$), we have

$$s_\lambda(1, q, \dots, q^{n-1}) = \frac{\det\left(q^{(i-1)(\lambda_j + n - j)}\right)_{i,j=1}^n}{\det\left(q^{(i-1)(n-j)}\right)_{i,j=1}^n}. \tag{7.104}$$

The denominator is just a specialization of the Vandermonde determinant $a_\delta = \det(x_i^{n-j}) = \prod_{1 \leq i < j \leq n}(x_i - x_j)$, and so is equal to $\prod_{1 \leq i < j \leq n}(q^{i-1} - q^{j-1})$. But the

numerator is also a specialization of a_δ, albeit somewhat disguised. Namely, let

$$a_\delta^* = \det\left(x_j^{i-1}\right)_{i,j=1}^n.$$

We have $a_\delta^* = (-1)^{\binom{n}{2}} a_\delta$ since the matrix $(x_j^{i-1})_{i,j=1}^n$ is obtained from that defining a_δ by transposing a_δ and then reversing the order of the rows. Thus the numerator of the right-hand side of (7.104) is just $a_\delta^*(q^{\mu_1}, q^{\mu_2}, \ldots, q^{\mu_n})$, where $\mu_j = \lambda_j + n - j$, so we get

$$s_\lambda(1, q, \ldots, q^{n-1}) = (-1)^{\binom{n}{2}} \prod_{1 \le i < j \le n} \frac{q^{\mu_i} - q^{\mu_j}}{q^{i-1} - q^{j-1}}. \tag{7.105}$$

By Lemma 7.21.1 there follows (using $\prod_{1\le i<j\le n}[j-i] = \prod_{i=1}^n [n-i]!$)

$$\begin{aligned} s_\lambda(1, q, \ldots, q^{n-1}) &= \frac{q^{\sum_{i<j}\mu_j} \prod_{i<j}[\mu_i - \mu_j] \cdot \prod_{i\ge 1}[\mu_i]!}{q^{\sum_{i<j}(i-1)} \prod_{i<j}[j-i] \cdot \prod_{i \ge 1}[\mu_i]!} \\ &= q^{b(\lambda)} \prod_{u \in \lambda} \frac{[n + c(u)]}{[h(u)]}. \end{aligned}$$

□

Note that

$$s_\lambda(1, q, \ldots, q^{n-1}) = \sum_\pi q^{|\pi|},$$

where π ranges over all column-strict plane partitions (= reverse SSYTs) of shape λ and largest part at most $n-1$, allowing 0 as a part. Hence Theorem 7.21.2 may be regarded as determining the generating function for this class of plane partitions, enumerated by the sum of their parts. If one prefers not to have 0 as a part, then the homogeneity of s_λ gives

$$\begin{aligned} q^{|\lambda|} s_\lambda(1, q, \ldots, q^{n-1}) &= s_\lambda(q, q^2, \ldots, q^n) \\ &= \sum_\pi q^{|\pi|}, \end{aligned}$$

where now π ranges over all column-strict plane partitions of shape λ and largest part at most n (with the usual condition that the parts are *positive* integers).

If we now let $n \to \infty$ in Theorem 7.21.2 then the numerator $\prod_{u\in\lambda}(1 - q^{n+c(u)})$ goes to 1, so we get a formula for the stable principal specialization $s_\lambda(1, q, q^2, \ldots)$.

7.21.3 Corollary. *For any* $\lambda \in \mathrm{Par}$ *we have*

$$s_\lambda(1, q, q^2, \ldots) = \frac{q^{b(\lambda)}}{\prod_{u\in\lambda}[h(u)]}.$$

Similarly, if we set $q = 1$ in Theorem 7.21.2, then we get (using the fact that $1 - q^k = (1-q)(1 + q + \cdots + q^{k-1})$ and canceling the factors of $1 - q$ from the numerator and denominator) the following result. (For its representation-theoretic significance, see Appendix 2, equation (A2.155) and the discussion following.)

7.21.4 Corollary. *For any $\lambda \in$ Par and $n \in \mathbb{P}$ we have*

$$s_\lambda(1^n) = \prod_{u \in \lambda} \frac{n + c(u)}{h(u)}. \qquad (7.106)$$

In particular, all the zeros of $s_\lambda(1^n)$ (regarded as a polynomial in n) are integers.

Corollary 7.21.3 has some interesting consequences. For instance, setting $\mu = \nu$ in Proposition 7.19.11 and comparing with Corollary 7.21.3 yields the following result.

7.21.5 Corollary. *For any $\lambda \in$ Par we have*

$$\sum_T q^{\mathrm{maj}(T)} = \frac{q^{b(\lambda)}[n]!}{\prod_{u \in \lambda}[h(u)]},$$

where T ranges over all SYTs of shape λ.

From Corollary 7.21.5 we obtain the explicit formula for f^λ mentioned after Corollary 7.16.3 and thus also an enumeration of the combinatorial objects given in Proposition 7.10.3. This remarkable result is known as the *hook-length formula*.

7.21.6 Corollary. *Let $\lambda \vdash n$. Then*

$$f^\lambda = \frac{n!}{\prod_{u \in \lambda} h(u)}.$$

Proof. Set $q = 1$ in Corollary 7.21.5. Alternatively, by equation (7.29) we can restate Corollary 7.21.3 for $\lambda \vdash n$ as

$$\mathrm{ex}_q(s_\lambda) = \frac{t^n q^{b(\lambda)}}{\prod_{u \in \lambda}\left(1 + q + \cdots + q^{h(u)-1}\right)}.$$

Now let $q = 1$ and use the interpretation of ex given by Proposition 7.8.4(a) □

We may regard both Corollaries 7.21.3 and 7.21.5 as q-analogues of the hook-length formula. Corollary 7.21.3 is the symmetric function q-analogue, while Corollary 7.21.5 is the combinatorial q-analogue.

Theorem 7.21.2 is a completely satisfactory generating function for column-strict plane partitions, but how is it related to ordinary plane partitions? The answer

is that for *rectangular shapes* $\lambda = \langle c^r \rangle$ there is a simple bijection between column-strict plane partitions of shape $\langle c^r \rangle$ and ordinary plane partitions of shape $\langle c^r \rangle$, and this bijection has an easily computable effect on the largest part and the sum of the parts. These assertions will be explained in the proof of the following theorem, which is the main result of this section.

7.21.7 Theorem. *Fix r, c, t with $r \le c$. Then*

$$\sum_{\pi \in B(r,c,t)} q^{|\pi|}$$

$$= \frac{[t+1][t+2]^2 \cdots [t+r]^r [t+r+1]^r \cdots [t+c]^r [t+c+1]^{r-1} \cdots [t+c+r-1]}{[1][2]^2 \cdots [r]^r [r+1]^r \cdots [c]^r [c+1]^{r-1} \cdots [c+r-1]}, \tag{7.107}$$

where $[i] = 1 - q^i$.

Proof. Let $\lambda = \langle c^r \rangle$, a rectangular shape with r rows and c columns. Note that the assumption $r \le c$ entails no loss of generality since we can always replace λ with λ'. Let $\pi = (\pi_{ij})$ be a column-strict plane partition of shape λ, allowing 0 as a part. Define $\pi^* = (\pi^*_{ij})$ by $\pi^*_{ij} = \pi_{ij} - r + i$. We have simply applied to each column of π the usual method of converting a strictly decreasing sequence into a weakly decreasing one. For instance, if

$$\begin{array}{rcccccc} & 6 & 6 & 4 & 4 & 4 & 3 \\ \pi = & 4 & 3 & 3 & 2 & 2 & 1 \\ & 2 & 2 & 1 & 1 & 0 & 0 \end{array}$$

then

$$\begin{array}{rcccccc} & 4 & 4 & 2 & 2 & 2 & 1 \\ \pi^* = & 3 & 2 & 2 & 1 & 1 & 0. \\ & 2 & 2 & 1 & 1 & 0 & 0 \end{array}$$

It is clear that π^* is a plane partition satisfying

$$\ell_1(\pi^*) \le r, \qquad \ell_2(\pi^*) \le c,$$
$$\max(\pi^*) = \max(\pi) - r + 1, \qquad |\pi^*| = |\pi| - \binom{r}{2} c.$$

Moreover, given such a plane partition π^*, we can recover π by $\pi_{ij} = \pi^*_{ij} + r - i$. Hence we obtain from Theorem 7.21.2 that

$$\sum_{\pi \subseteq B(r,c,t)} q^{|\pi|} = q^{-\binom{r}{2}c} s_{\langle c^r \rangle}(1, q, \ldots, q^{t+r-1})$$
$$= q^{b(\langle c^r \rangle) - \binom{r}{2}c} \prod_{u \in \langle c^r \rangle} \frac{[t+r+c(u)]}{[h(u)]}. \tag{7.108}$$

Note that $b(\langle c^r\rangle) = \binom{r}{2}c$. Moreover, the multiset of hook lengths of $\langle c^r\rangle$ is $\{1, 2^2, 3^3, \ldots, r^r, (r+1)^r, \ldots, c^r, (c+1)^{r-1}, (c+2)^{r-2}, \ldots, c+r-1\}$, and the multiset of contents is obtained by subtracting r from the hook lengths. Substituting these values of $c(u)$ and $h(u)$ into (7.108) completes the proof. □

The reader can check that an alternative way of writing the generating function (7.107) that shows more clearly the symmetry between r, c, and t (but that has rct factors in the numerator and denominator rather than rc factors) is given by

$$\sum_{\pi\in B(r,c,t)} q^{|\pi|} = \prod_{i=1}^{r}\prod_{j=1}^{c}\prod_{k=1}^{t} \frac{[i+j+k-1]}{[i+j+k-2]}. \tag{7.109}$$

Theorem 7.21.7 can be interpreted in terms of the theory of P-partitions developed in Section 4.5. A plane partition π satisfying $\ell_1(\pi) \le r$, $\ell_2(\pi) \le c$, and $\max(\pi) \le t$ may be regarded as an order-reversing map $\pi : \mathbf{r}\times\mathbf{c} \to [0, t]$. In other words, if P is the poset $\mathbf{r}\times\mathbf{c}$, then π is just a P-partition with largest part at most t. Thus in particular the number of such π is just $\Omega(P, t+1)$, where Ω denotes the order polynomial. Setting $q = 1$ and $t = m - 1$ in (7.107) yields

$$\Omega(\mathbf{r}\times\mathbf{c}, m) = \frac{m(m+1)^2\cdots(m+r-1)^r(m+r)^r\cdots(m+c-1)^r(m+c)^{r-1}\cdots(m+c+r-2)}{1\cdot 2^2\cdots r^r(r+1)^r\cdots c^r(c+1)^{r-1}\cdots(c+r-1)}.$$

In particular, all the zeros of $\Omega(\mathbf{r}\times\mathbf{c}, m)$ are nonpositive integers. Moreover, a simple extension of Proposition 3.5.1 (see Exercise 4.24(a)) shows that

$$\sum_{\pi\in B(r,c,t)} q^{|\pi|} = \sum_{I\in J(\mathbf{r}\times\mathbf{c}\times\mathbf{t})} q^{\#I}, \tag{7.110}$$

the rank-generating function of the distributive lattice $J(\mathbf{r}\times\mathbf{c}\times\mathbf{t})$. Thus equation (7.109) is equivalent to Exercises 3.27(b) and 4.25(f)(i). Setting $q = 1$ in (7.110) and (7.109) yields the elegant formula

$$\#J(\mathbf{r}\times\mathbf{c}\times\mathbf{t}) = \prod_{i=1}^{r}\prod_{j=1}^{c}\prod_{k=1}^{t} \frac{i+j+k-1}{i+j+k-2}.$$

7.22 Reverse Plane Partitions and the Hillman–Grassl Correspondence

The proof of Theorem 7.21.7 involved a bijection between column-strict plane partitions of shape $\langle c^r\rangle$ and ordinary plane partitions whose shape is contained in $\langle c^r\rangle$. It is natural to ask whether we can do a similar bijection for any shape λ,

thereby obtaining a formula for the generating function

$$\sum_{\substack{\mathrm{sh}(\pi)\subseteq\lambda\\ \max(\pi)\le t}} q^{|\pi|}, \tag{7.111}$$

where π ranges over all plane partitions whose shape is contained in λ and with largest part $\le t$. Unfortunately the bijection used in the proof of Theorem 7.21.7 does not carry over to nonrectangular shapes. For instance, if $\lambda = (2, 1)$ then we could associate with the ordinary plane partition (where 0 is allowed as a part)

$$\pi = \begin{matrix} a & b \\ c & \end{matrix}$$

the column-strict plane partition (also allowing 0 as a part)

$$\pi' = \begin{matrix} a+1 & b+1 \\ c & \end{matrix}.$$

But then there is no π corresponding to $\pi' = \begin{matrix}10\\0\end{matrix}$. Similarly, if we instead tried

$$\pi' = \begin{matrix} a+1 & b \\ c & \end{matrix},$$

then there is no π corresponding to $\pi' = \begin{matrix}11\\0\end{matrix}$ (since $\begin{matrix}01\\0\end{matrix}$ is not a plane partition). Indeed, the generating function (7.111), even in the case $t = \infty$, does not in general factor into a simple product (though there does exist a determinantal formula for (7.111)).

Although there is no such simple correspondence between column-strict plane partitions and plane partitions of a given shape as was used to prove Theorem 7.21.7, such a correspondence does exist in the *reverse* situation. This correspondence does not have a uniform effect on the largest part, but it is well behaved with respect to the sum of the parts. Thus we will get the generating function for reverse plane partitions of n of shape λ. It is easier to work with *weak* reverse plane partitions of shape λ, for which 0 can be a part. To get from a weak reverse plane partition of shape λ to a (nonweak) reverse plane partition λ^* of shape λ, simply add 1 to every entry of π. Note that $|\pi^*| = |\pi| + |\lambda|$. Similary define a *weak* SSYT to be an SSYT in which 0 is allowed to be a part.

7.22.1 Theorem. *Let* $\lambda \in$ Par. *Then*

$$\sum_{\pi} q^{|\pi|} = \frac{1}{\prod_{u\in\lambda}[h(u)]}, \tag{7.112}$$

where π ranges over all weak reverse plane partitions of shape λ. (If one does not want to allow 0 *as a part, simply multiply by* $q^{|\lambda|}$.)

Proof. Let $T = (a_{ij})$ be a weak SSYT of shape λ. Define $\pi = \pi(T)$ by $\pi_{ij} = a_{ij} - i + 1$. Then π is a weak reverse plane partition of shape λ satisfying

$$|\pi| = |T| - \sum (i-1)\lambda_i = |T| - b(\lambda).$$

This correspondence is easily seen to be a bijection. Hence if $\mathcal{R}_\lambda$ (respectively, $\mathcal{R}'_\lambda$) denotes the set of all weak reverse plane partitions (respectively, weak SSYTs) of shape λ, then

$$\begin{aligned}\sum_{\pi\in\mathcal{R}_\lambda} q^{|\pi|} &= q^{-b(\lambda)} \sum_{T\in\mathcal{R}'_\lambda} q^{|T|}\\ &= q^{-b(\lambda)} s_\lambda(1, q, q^2, \ldots),\end{aligned}$$

and the proof follows from Corollary 7.21.3. □

Theorem 7.22.1 is so elegant that we could ask for a simple bijective proof. Identifying λ with its diagram, we want a bijection between weak reverse plane partitions π of shape λ and functions $f : \lambda \to \mathbb{N}$, such that

$$|\pi| = \sum_{u\in\lambda} f(u)h(u).$$

We now describe such a bijection, known as the *Hillman–Grassl correspondence*.

We will successively define pairs $(\pi_0, f_0), (\pi_1, f_1), \ldots, (\pi_k, f_k)$, where each π_i is a weak reverse plane partition of shape λ and $f_i : \lambda \to \mathbb{N}$. We begin with $\pi_0 = \pi$ and $f_0(u) = 0$ for all $u \in \lambda$. We obtain π_{i+1} from π_i by decreasing $h(u_i)$ of the entries of π_i by 1 for a certain square $u_i \in \lambda$ (to be explained), and we define

$$f_{i+1}(v) = \begin{cases} f_i(v), & v \neq u_i \\ f_i(u_i) + 1, & v = u_i. \end{cases}$$

At the end π_k will have every entry equal to 0. Consequently,

$$|\pi| = \sum_{u\in\lambda} f_k(u)h(u),$$

and we define $f = f_k$.

It remains to describe the rule for obtaining π_{i+1} from π_i, and the corresponding choice of $u_i \in \lambda$. We will define a lattice path L in λ with steps N or E (i.e., one square up or one square to the right), beginning at the bottom of a column and ending at the end of a row of λ. The lattice path L begins at the location of the

southwesternmost nonzero entry of π_i. If the path has reached the square (a, b), then move N if $(\pi_i)_{ab} = (\pi_i)_{a-1,b} > 0$; otherwise if $(\pi_i)_{ab} > 0$ move E. The lattice path terminates when no further move is possible. Define π_{i+1} to be the array obtained from π_i by subtracting 1 from every entry that lies in a square of the path L. If L begins in column b and ends in row a, then let $u_i = (a, b)$. It is easy to see that $\#L = h(u_i)$.

We illustrate this correspondence with a reverse plane partition π of shape $(3, 3, 1)$. We indicate the lattice path L in boldface, and the function f_i by putting the value $f_i(u)$ in the square $u \in \lambda$:

π_i	f_i
0 1 3	0 0 0
2 4 4	0 0 0
3	0
0 1 3	0 0 0
2 4 4	0 0 0
2	1
0 1 **3**	0 0 0
1 3 3	1 0 0
1	1
0 1 **2**	1 0 0
0 **2 2**	1 0 0
0	1
0 **1 1**	1 1 0
0 **1** 1	1 0 0
0	1
0 0 0	1 2 0
0 0 **1**	1 0 0
0	1
0 0 0	1 2 0
0 0 0	1 0 1
0	1

We omit the proof that this correspondence is indeed a bijection, except for the hint that the square u_i is in the rightmost column among all the squares $u_1, \dots, u_i$, and is in the highest row among the squares $u_1, \dots, u_i$ in its column.

Now let P be an n-element poset, and let $\mathcal{A}(\mathcal{P})$ denote the set of P-partitions as in Section 4.5. Recall from Theorem 4.5.6 that

$$G_P(q) := \sum_{\sigma \in \mathcal{A}(\mathcal{P})} q^{|\sigma|} = \frac{W_P(q)}{(1-q)\cdots(1-q^n)},$$

where $W_P(1) = e(P)$, the number of linear extensions of P. If we set $P = P_\lambda^*$, the dual of the poset P_λ defined after Corollary 7.19.5, then the left-hand side of

equation (7.112) is just $G_P(q)$. Thus we get

$$f^\lambda = e(P_\lambda) = e\big(P_\lambda^*\big) = \left.\frac{[n]!}{\prod_{u\in\lambda}[h(u)]}\right|_{q=1} = \frac{n!}{\prod_{u\in\lambda}h(u)} \tag{7.113}$$

giving a proof of the hook-length formula (Corollary 7.21.6) avoiding the use of determinants.

7.23 Applications to Permutation Enumeration

We have already seen a number of connections between symmetric functions and permutations (such as the RSK algorithm applied to a permutation matrix, or the Jordan–Hölder set of the labeled poset $(P_{\lambda/\mu}, \omega_{\lambda/\mu})$), so it is not too surprising that the theory of symmetric functions can be used to obtain some results related to permutation enumeration. Our first result is based on the RSK algorithm, and requires one additional fact about it.

7.23.1 Lemma. *Let $w \in \mathfrak{S}_n$ and $w \xrightarrow{\text{RSK}} (P, Q)$. Then $D(P) = D(w^{-1})$ and $D(Q) = D(w)$, where D denotes the descent set.*

Proof. Let $(P_0, Q_0), \ldots, (P_n, Q_n) = (P, Q)$ be the successive pairs of tableaux obtained in applying the RSK algorithm to w. Let $w = w_1 \cdots w_n$, and suppose that for some i we have $w_i < w_{i+1}$. As observed in the proof of Theorem 7.11.5, the insertion path of w_{i+1} lies to the right of that of w_i. Suppose that the shape of P_i is obtained from that of P_{i-1} by adjoining a square in the (a, b) position, so the (a, b) entry of Q is equal to i. When we insert w_{i+1} into P_i, if an element m is bumped into row a then it would occupy the $(a, b+1)$ position without bumping a further element. Thus $i+1$ does not appear in Q in a lower row than i, so $i \notin D(Q)$.

Similarly if $w_i > w_{i+1}$ then the insertion path of w_{i+1} lies weakly to the left of that of w_i. Thus an element must be bumped into row a but not at the end, and hence must bump an element into row $a+1$. This means $i \in D(Q)$, so $D(w) = D(Q)$ as claimed.

The symmetry property Theorem 7.13.1 of the RSK algorithm implies that $w^{-1} \xrightarrow{\text{RSK}} (Q, P)$, so by what was just proved we have $D(w^{-1}) = P$, completing the proof. □

Recall from Section 7.19 the definition $\mathrm{co}(S)$ of a set $S \subseteq [n-1]$, and the associated definitions of $\mathrm{co}(w)$ for $w \in \mathfrak{S}_n$ and $\mathrm{co}(T)$ when T is an SYT. Similarly define $\mathrm{co}'(w) = \mathrm{co}([n-1] - D(w))$. Note that $[n-1] - D(w)$ is just the ascent set $A(w)$ of w.

As a corollary to Lemma 7.23.1, we obtain the expansion of the Cauchy product $\prod(1 - x_i y_j)^{-1}$ in terms of quasisymmetric functions. (We regard n as fixed, so all our quasisymmetric functions are of degree n.) This result may be regarded as giving a generating function for the number of permutations $w \in \mathfrak{S}_n$ such that $D(w^{-1}) = S$ and $D(w) = T$, since this number is just the coefficient of $L_\alpha(x)L_\beta(y)$ in the expansion below, where $\alpha = \alpha_S$ and $\beta = \beta_T$.

7.23.2 Theorem. *Let $n \in \mathbb{N}$. Then*

$$\sum_{\lambda \vdash n} s_\lambda(x) s_\lambda(y) = \sum_{w \in \mathfrak{S}_n} L_{\mathrm{co}(w^{-1})}(x) L_{\mathrm{co}(w)}(y) \tag{7.114}$$

$$\sum_{\lambda \vdash n} s_\lambda(x) s_{\lambda'}(y) = \sum_{w \in \mathfrak{S}_n} L_{\mathrm{co}(w^{-1})}(x) L_{\mathrm{co}'(w)}(y). \tag{7.115}$$

Proof. By the quasisymmetric expansion of s_λ (Theorem 7.19.7), we have

$$\begin{aligned}\sum_{\lambda \vdash n} s_\lambda(x) s_\lambda(y) &= \sum_{\lambda \vdash n} \left(\sum_{\mathrm{sh}(T)=\lambda} L_{\mathrm{co}(T)}(x) \right) \left(\sum_{\mathrm{sh}(T')=\lambda} L_{\mathrm{co}(T')}(y) \right) \\ &= \sum_{\lambda \vdash n} \sum_{\substack{\mathrm{sh}(T)=\lambda \\ \mathrm{sh}(T')=\lambda}} L_{\mathrm{co}(T)}(x) L_{\mathrm{co}(T')}(y),\end{aligned}$$

where $\mathrm{sh}(T) = \lambda$ signifies that the sum ranges over all SYT of shape λ (and similarly for $\mathrm{sh}(T') = \lambda$). If $w \in \mathfrak{S}_n$ satisfies $w \xrightarrow{\mathrm{RSK}} (T, T')$, then by Lemma 7.23.1 we have $D(w) = T'$ and $D(w^{-1}) = T$. Hence

$$\sum_{\lambda \vdash n} \sum_{\substack{\mathrm{sh}(T)=\lambda \\ \mathrm{sh}(T')=\lambda}} L_{\mathrm{co}(T)}(x) L_{\mathrm{co}(T')}(y) = \sum_{w \in \mathfrak{S}_n} L_{\mathrm{co}(w^{-1})}(x) L_{\mathrm{co}(w)}(y),$$

and (7.114) follows. The proof of (7.115) is analogous, using the dual RSK algorithm. Alternatively, apply the extension $\hat{\omega}$ of ω (acting on the y variables only) given by Exercise 7.94(a) to (7.114). □

Although Theorem 7.23.2 may be regarded as "determining" the number of permutations $w \in \mathfrak{S}_n$ with $D(w^{-1}) = S$ and $D(w) = T$, a more useful or explicit expression would be desirable. Such an expression can be given in terms of skew Schur functions whose shape is a border strip. If $\alpha = (\alpha_1, \ldots, \alpha_\ell) \in \mathrm{Comp}(n)$, then let B_α denote the border strip with α_i squares in row $\ell - i + 1$. Regard B_α as a skew shape, so s_{B_α} is a skew Schur function.

7.23.3 Lemma. *The Jordan–Hölder set $\mathcal{L}(P_{B_\alpha}, \omega_{B_\alpha})$ (as defined in Section 7.19) consists of all permutations $w \in \mathfrak{S}_n$ satisfying $\mathrm{co}(w^{-1}) = \alpha$.*

Before proceeding to the proof, let us consider an example. Let $\alpha = (3, 1, 2, 3) \in \mathrm{Comp}(9)$. The corresponding Schur labeled border strip is

$$\begin{array}{cccccc} & & & 7 & 8 & 9 \\ & & 5 & 6 & & \\ & & 4 & & & \\ 1 & 2 & 3 & & & \end{array}$$

A typical element of $\mathcal{L}(P,\omega)$ is $w = 578124963$. Then $w^{-1} = 459618237$, so $D(w^{-1}) = \{3,4,6\}$ and $\mathrm{co}(w^{-1}) = (3,1,2,3) = \alpha$.

Proof of Lemma 7.23.3. A permutation $w \in \mathfrak{S}_n$ belongs to $\mathcal{L}(P_{B_\alpha}, \omega_{B_\alpha})$ if and only if $i+1$ follows i in w (regarded as a word $w_1 \cdots w_n$) whenever i and $i+1$ are in the same row, and $i+1$ precedes i in w whenever i and $i+1$ are in the same column. Hence $i \in D(w^{-1})$ if and only if i and $i+1$ are in the same column, which is clearly equivalent to $i \in S_\alpha$. □

7.23.4 Corollary. *Let $\alpha \in \mathrm{Comp}(n)$. Then*

$$s_{B_\alpha} = \sum_{\substack{w \in \mathfrak{S}_n \\ \alpha = \mathrm{co}(w^{-1})}} L_{\mathrm{co}(w)}.$$

Proof. Immediate from Theorem 7.19.7 and Lemma 7.23.3. □

7.23.5 Corollary. *Let $n \in \mathbb{N}$. Then*

$$\sum_{\lambda \vdash n} s_\lambda(x) s_\lambda(y) = \sum_{\alpha \in \mathrm{Comp}(n)} L_\alpha(x) s_{B_\alpha}(y). \tag{7.116}$$

Proof. By Corollary 7.23.4 we have

$$\begin{aligned} \sum_{\alpha \in \mathrm{Comp}(n)} L_\alpha(x) s_{B_\alpha}(y) &= \sum_{\alpha \in \mathrm{Comp}(n)} L_\alpha(x) \sum_{\substack{w \in \mathfrak{S}_n \\ \mathrm{co}(w^{-1}) = \alpha}} L_{\mathrm{co}(w)}(y) \\ &= \sum_{w \in \mathfrak{S}_n} L_{\mathrm{co}(w^{-1})}(x) L_{\mathrm{co}(w)}(y), \end{aligned}$$

and the proof follows from Theorem 7.23.2. □

The next corollary gives a formula for the expansion of any symmetric function in terms of fundamental quasisymmetric functions.

7.23.6 Corollary. *For all $f \in \Lambda^n$ we have*

$$f = \sum_{\alpha \in \mathrm{Comp}(n)} \langle f, s_{B_\alpha} \rangle L_\alpha.$$

Proof. Take the scalar product of both sides of equation (7.116) with $s_\lambda(y)$ to obtain the desired result for $f = s_\lambda$. The general case follows by linearity. □

Next we come to an alternative expansion of $\sum_{\lambda \vdash n} s_\lambda(x) s_\lambda(y)$ in terms of quasisymmetric functions.

7.23.7 Corollary. *Let $n \in \mathbb{N}$. Then*

$$\sum_{\lambda \vdash n} s_\lambda(x)s_\lambda(y) = \sum_{\alpha,\beta \in \mathrm{Comp}(n)} \langle s_{B_\alpha}, s_{B_\beta}\rangle L_\alpha(x)L_\beta(y).$$

Proof. By Corollary 7.23.6 we have

$$s_{B_\beta}(y) = \sum_{\alpha \in \mathrm{Comp}(n)} \langle s_{B_\alpha}, s_{B_\beta}\rangle L_\alpha.$$

Substitute into the right-hand side of equation (7.116) to complete the proof. □

Write B_S and B_T for the border strips B_{α_S} and B_{α_T}, where α_S and α_T are the compositions corresponding to S and T as defined in Section 7.19. Comparing Theorem 7.23.2 with Corollary 7.23.7 yields the following enumerative result.

7.23.8 Corollary. *Let $S, T \subseteq [n-1]$. Then the number of permutations $w \in \mathfrak{S}_n$ satisfying $D(w^{-1}) = S$ and $D(w) = T$ is equal to the scalar product $\langle s_{B_S}, s_{B_T}\rangle$.*

Theorem 7.23.2 can be specialized in several ways. For instance, we can obtain a generating function for the joint distribution of the statistics $\mathrm{maj}(w)$ and $\mathrm{maj}(w^{-1})$ for $w \in \mathfrak{S}_n$. Write $[k]_q = 1 - q^k$, $[k]_t = 1 - t^k$, $[k]!_q = [1]_q \cdots [k]_q$, and $[k]!_t = [1]_t \cdots [k]_t$. Recall the notation $b(\lambda) = \sum_i (i-1)\lambda_i$ from equation (7.103).

7.23.9 Corollary. *Let*

$$F_n(q,t) = \sum_{w \in \mathfrak{S}_n} q^{\mathrm{maj}(w^{-1})} t^{\mathrm{maj}(w)}.$$

Then

$$F_n(q,t) = \sum_{\lambda \vdash n} \frac{q^{b(\lambda)}\, t^{b(\lambda)}\, [n]!_q\, [n]!_t}{\prod_{u \in \lambda} [h(u)]_q\, [h(u)]_t}, \tag{7.117}$$

and we have the "double Eulerian" generating function

$$\begin{aligned}\sum_{n \geq 0} F_n(q,t) \frac{z^n}{[n]!_q\, [n]!_t} &= \prod_{i,j \geq 0} \frac{1}{1 - q^i t^j z} \\ &= \exp \sum_{n \geq 1} \frac{1}{n} \frac{z^n}{(1-q^n)(1-t^n)}.\end{aligned}$$

Proof. Let $x_i = q^{i-1}$ and $y_i = t^{i-1}$ in equation (7.114). By Corollary 7.21.3, the left-hand side of (7.114) becomes

$$\sum_{\lambda \vdash n} s_\lambda(1,q,q^2,\ldots) s_\lambda(1,t,t^2,\ldots) = \sum_{\lambda \vdash n} \frac{q^{b(\lambda)}\, t^{b(\lambda)}}{\prod_{u \in \lambda} [h(u)]_q\, [h(u)]_t}.$$

On the other hand, by Lemma 7.19.10 the right-hand side of (7.114) becomes

$$\sum_{w\in\mathfrak{S}_n} L_{\mathrm{co}(w^{-1})}(1,q,q^2,\ldots)L_{\mathrm{co}(w)}(1,t,t^2,\ldots)=\sum_{w\in\mathfrak{S}_n}\frac{q^{\mathrm{comaj}(w^{-1})}\,t^{\mathrm{comaj}(w)}}{[n]!_q\,[n]!_t},$$

where $\mathrm{comaj}(w)=\sum_{i\in D(w)}(n-i)$. If $w=w_1\cdots w_n$ then define $w^*=n+1-w_n,\ldots,n+1-w_1$. The map $w\mapsto w^*$ from $\mathfrak{S}_n$ to itself is a bijection (in fact, involution), satisfying $\mathrm{comaj}(w)=\mathrm{maj}(w^*)$ and $(w^{-1})^*=(w^*)^{-1}$. Hence

$$\sum_{w\in\mathfrak{S}_n} q^{\mathrm{comaj}(w^{-1})}t^{\mathrm{comaj}(w)}=\sum_{w\in\mathfrak{S}_n} q^{\mathrm{maj}(w^{-1})}t^{\mathrm{maj}(w)},$$

so equation (7.117) follows. The remainder of the proof is an immediate consequence of Proposition 7.7.4 and Theorem 7.12.1, which assert that

$$\sum_{\lambda\in\mathrm{Par}} s_\lambda(x)s_\lambda(y)=\prod(1-x_iy_j)^{-1}=\exp\sum_{n\geq 1}\frac{1}{n}p_n(x)p_n(y). \qquad \square$$

It is not necessary to use the theory of symmetric functions to prove Corollary 7.23.9; see for instance Exercise 4.20 (in the case $m=2$).

Our second connection between symmetric functions (more accurately, the RSK algorithm) and permutation enumeration concerns increasing and decreasing subsequences of a permutation. If $w=w_1\cdots w_n\in\mathfrak{S}_n$, then let $v=w_{i_1}w_{i_2}\cdots w_{i_k}$ be a *subsequence* of w, i.e., $1\leq i_1<i_2<\cdots<i_k\leq n$. We say v is *increasing* if $w_{i_1}<w_{i_2}<\cdots<w_{i_k}$ and *decreasing* if $w_{i_1}>w_{i_2}>\cdots>w_{i_k}$. Write $\mathrm{is}(w)$ for the length (number of terms) of the longest increasing subsequence of w. Let $r_i(w)$ be the rightmost integer j in w such that the longest increasing subsequence of w whose last term is j has length i. If, for instance, $w=725481963$, then $\mathrm{is}(w)=4$, $r_1(w)=1$, $r_2(w)=3$, $r_3(w)=6$, and $r_4(w)=9$, while $r_i(w)$ is undefined for $i>4$. Note that in general $1=r_1(w)<r_2(w)<\cdots<r_{\mathrm{is}(w)}(w)$.

7.23.10 Proposition. *Let $w\in\mathfrak{S}_n$ and $m=\mathrm{is}(w)$. Suppose that $w\xrightarrow{\mathrm{RSK}}(P,Q)$. Then the first row of P is equal to $r_1(w),r_2(w),\ldots,r_m(w)$.*

Proof. The proposition is essentially a restatement of Lemma 7.13.4. An increasing subsequence $w_{i_1},\ldots,w_{i_k}$ of w is equivalent to a chain $(i_1,w_{i_1})<\cdots<(i_k,w_{i_k})$ in the inversion poset $I(w)$. If follows that the antichain I_j consists precisely of those pairs (i,w_i) for which the longest increasing subsequence of w ending at w_i has length j. The maximum value of i for such a pair is by definition u_{jn_j}, and the corresponding value of w_i is equal to v_{jn_j}. Hence $v_{jn_j}=r_j(w)$, so the proof follows from Lemma 7.13.4. $\square$

As an immediate corollary, we obtain a combinatorial interpretation of the length of the first row of P (or Q) when $w\xrightarrow{\mathrm{RSK}}(P,Q)$.

7.23.11 Corollary. *Suppose that* $w \in \mathfrak{S}_n$ *and* $w \xrightarrow{\text{RSK}} (P, Q)$. *Let* $\mathrm{sh}(P) = \mathrm{sh}(Q) = \lambda$. *Then* $\lambda_1 = \mathrm{is}(w)$.

Corollary 7.23.11 can be used to obtain interesting enumerative results concerning the distribution of longest increasing subsequences. A quite general result is the following, though it is often difficult to extract further information from it. (See, for instance, Exercise 6.56 and Exercise 7.16.)

7.23.12 Corollary. *Let* $g_p(n)$ *denote the number of permutations* $w \in \mathfrak{S}_n$ *for which* $\mathrm{is}(w) = p$. *Then*

$$g_p(n) = \sum_{\substack{\lambda \vdash n \\ \lambda_1 = p}} (f^\lambda)^2.$$

Proof. There are $(f^\lambda)^2$ pairs (P, Q) of SYTs of shape λ. The proof now follows from Corollary 7.23.11. □

If $w \in \mathfrak{S}_n$, let us define the *shape* $\mathrm{sh}(w)$ to be the shape of the SYT P or Q when $w \xrightarrow{\text{RSK}} (P, Q)$. If $\lambda = \mathrm{sh}(w)$ then we have found a simple combinatorial interpretation of the largest part λ_1 of λ. It is natural to ask for a similar interpretation of the other parts λ_i of λ. For instance, it is tempting to conjecture that λ_2 is equal to the length of the longest possible increasing subsequence that can remain when an increasing subsequence of length λ_1 is removed from w. Unfortunately this conjecture is false. For instance, if $w = 247951368$ then $\mathrm{sh}(w) = (5, 3, 1)$. There is a unique increasing subsequence of w of length 5, viz., 24568. When this is removed from w, we obtain the sequence 7913, which has no increasing subsequence of length 3. The correct result is given by the following fundamental theorem, whose proof is included in Appendix 1 (Theorem A1.1.1).

7.23.13 Theorem. *Let* $w \in \mathfrak{S}_n$ *and* $\mathrm{sh}(w) = (\lambda_1, \lambda_2, \ldots)$. *Then for all* $i \geq 1$, $\lambda_1 + \cdots + \lambda_i$ *is equal to the length of the longest subsequence of* w *that can be written as a union of* i *increasing subsequences.*

For instance, let $w = 247951368$ as above. The subsequence 24791368 is the union of the two increasing subsequences 2479 and 1368. Hence $\lambda_1 + \lambda_2 \geq 8$. In fact w itself cannot be written as a union of two increasing subsequences, so actually $\lambda_1 + \lambda_2 = 8$.

Instead of increasing subsequences we can ask about both increasing and decreasing subsequences simultaneously. The key to this question is a further symmetry property of the RSK algorithm (in addition to Theorem 7.13.1). We outline one approach to this symmetry property here, while a different method of proof is given in Appendix 1 (Corollary A1.2.11). We have denoted the row insertion of

the integer k into the SSYT T by $T \leftarrow k$. Assume that all entries of T are distinct and different from k, and let $k \rightarrow T$ denote the *column* insertion of k into T. This is defined exactly like row insertion, but with the roles of rows and columns interchanged. Equivalently, if ^{t} denotes transpose then

$$(k \rightarrow T) = (T^t \leftarrow k)^t.$$

We omit the proof, which consists of an elementary but tedious analysis of cases, of the following fundamental lemma. We are always assuming that our tableaux have distinct entries, and that this condition is maintained after inserting further elements.

7.23.14 Lemma. *If $i \neq j$ then*

$$j \rightarrow (T \leftarrow i) = (j \rightarrow T) \leftarrow i.$$

In other words, row insertion and column insertion commute with each other.

7.23.15 Lemma. *Let*

$$\begin{aligned} P(i_1, i_2, \ldots, i_n) &= ((i_1 \leftarrow i_2) \leftarrow i_3) \leftarrow \cdots \leftarrow i_n \\ \tilde{P}(i_1, i_2, \ldots, i_n) &= i_1 \rightarrow \cdots \rightarrow (i_{n-2} \rightarrow (i_{n-1} \rightarrow i_n)). \end{aligned}$$

Then

$$P(i_1, i_2, \ldots, i_n) = \tilde{P}(i_1, i_2, \ldots, i_n).$$

Proof. Induction on n. The assertion is clear for $n = 1$, since $P(i_1) = \tilde{P}(i_1) = i_1$. It is also easy to check directly the case $n = 2$. Now let $n \geq 2$ and assume the assertion for all $m \leq n$. We have

$$\begin{aligned} P(i_1, \ldots, i_{n+1}) &= P(i_1, \ldots, i_n) \leftarrow i_{n+1} && \text{(definition of } \leftarrow) \\ &= \tilde{P}(i_1, \ldots, i_n) \leftarrow i_{n+1} && \text{(induction hypothesis)} \\ &= [i_1 \rightarrow \tilde{P}(i_2, \ldots, i_n)] \leftarrow i_{n+1} && \text{(definition of } \tilde{P}) \\ &= i_1 \rightarrow [\tilde{P}(i_2, \ldots, i_n) \leftarrow i_{n+1}] && \text{(previous lemma)} \\ &= i_1 \rightarrow [P(i_2, \ldots, i_n) \leftarrow i_{n+1}] && \text{(induction hypothesis)} \\ &= i_1 \rightarrow P(i_2, \ldots, i_n, i_{n+1}) && \text{(definition of } \leftarrow) \\ &= i_1 \rightarrow \tilde{P}(i_2, \ldots, i_n, i_{n+1}) && \text{(induction hypothesis)} \\ &= \tilde{P}(i_1, \ldots, i_{n+1}) && \text{(definition of } \rightarrow). \end{aligned}$$ □

We now come to the new symmetry property of the RSK algorithm mentioned above.

7.23.16 Theorem. *Let $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$ and $w \xrightarrow{\text{RSK}} (P, Q)$. Let $w^r = w_n \cdots w_2 w_1$, the word w written in reverse order. Suppose that $w^r \xrightarrow{\text{RSK}} (P^*, Q^*)$.*

Then $P^* = P^t$, *the transpose of* P. *In particular,* $\mathrm{sh}(w) = \mathrm{sh}(w^r)'$. *(The description of* Q^* *is more complicated and is discussed in Appendix 1, Section A1.2. The map* $Q \mapsto Q^*$ *is called the* Schützenberger involution.*)*

Proof. Using the notation of the previous lemma, we have $P(w_1, \ldots, w_n) = P$ and $\tilde{P}(w_1, \ldots, w_n)^t = P^*$. The proof follows from the previous lemma. $\square$

If we regard a permutation $w \in \mathfrak{S}_n$ as an $n \times n$ permutation matrix, then Theorem 7.13.1 tells us the effect on the RSK algorithm of reflecting w about the main diagonal, while Theorem 7.23.16 tells us the effect on the RSK algorithm of reflecting w about a horizontal line. These two reflections generate the entire eight-element dihedral group D_4 of symmetries of the square. Thus every "dihedral symmetry" of w has a predictable effect on the behavior of the RSK algorithm (when applied to a permutation).

Since a decreasing subsequence of w becomes an increasing subsequence (in reverse order) of w^r and vice versa, the following result is an immediate consequence of Theorem 7.23.13 and Theorem 7.23.16.

7.23.17 Theorem. *Let* $w \in \mathfrak{S}_n$ *and* $\mathrm{sh}(w) = \lambda$. *Then for all* $i \geq 1$, $\lambda'_1 + \cdots + \lambda'_i$ *is equal to the length of the longest subsequence of* w *that can be written as a union of* i *decreasing subsequences. In particular,* λ'_1 *is the length of the longest decreasing subsequence of* w.

Write $\mathrm{ds}(w)$ for the length of the longest decreasing subsequence of w. In the same way that Corollary 7.23.12 was obtained from Proposition 7.23.10, we deduce from both Proposition 7.23.10 and Theorem 7.23.17 the following result.

7.23.18 Corollary. *Let* $g_{p,q}(n)$ *denote the number of permutations* $w \in \mathfrak{S}_n$ *satisfying* $\mathrm{is}(w) = p$ *and* $\mathrm{ds}(w) = q$. *Then*

$$g_{p,q}(n) = \sum_{\substack{\lambda \vdash n \\ \lambda_1 = p,\ \lambda'_1 = q}} (f^\lambda)^2.$$

7.23.19 Example. (a) If $w \in \mathfrak{S}_{pq+1}$ then either $\mathrm{is}(w) > p$ or $\mathrm{ds}(w) > q$, since no partition $\lambda \vdash pq + 1$ satisfies $\lambda_1 \leq p$ and $\lambda'_1 \leq q$.

(b) Exactly one partition $\lambda \vdash pq$ satisfies $\lambda_1 = p$ and $\lambda'_1 = q$, viz., $\lambda = \langle p^q \rangle$. Hence, assuming $p \leq q$ (which entails no real loss of generality, since $g_{p,q}(n) = g_{q,p}(n)$), we get from Corollary 7.23.18 and the hook-length formula (Corollary 7.21.6) that

$$\begin{aligned} g_{p,q}(pq) &= \left(f^{\langle p^q \rangle}\right)^2 \\ &= \left(\frac{(pq)!}{1^1 2^2 \cdots p^p (p+1)^p \cdots q^p (q+1)^{p-1} \cdots (p+q-1)}\right)^2. \end{aligned}$$

(c) Let $p, q > n$. There are exactly $p(n)$ partitions (where $p(n)$ is the number of partitions of n) $\lambda \vdash p+q+n-1$ satisfying $\lambda_1 = p$ and $\lambda_1' = q$, viz., $\lambda = (p, 1+\mu_1, 1+\mu_2, \dots, 1+\mu_{q-1})$ where $\mu \vdash n$. Write for short $\lambda = (p, 1+\mu)$. Then

$$g_{p,q}(p+q+n-1) = \sum_{\mu \vdash n} \left(f^{(p,1+\mu)}\right)^2.$$

We can combine information concerning descent sets together with increasing and decreasing subsequences. For instance, the following result should be apparent to any reader who has followed this section up to here.

7.23.20 Proposition. *Let $g_{p,q,S,T}(n)$ denote the number of permutations $w \in \mathfrak{S}_n$ satisfying $\mathrm{is}(w) = p$, $\mathrm{ds}(w) = q$, $D(w^{-1}) = S$, $D(w) = T$. Then for fixed p, q, and n,*

$$\sum_{S,T} g_{p,q,S,T}(n) L_{\alpha_S}(x) L_{\alpha_T}(y) = \sum_{\substack{\lambda \vdash n \\ \lambda_1 = p,\ \lambda_1' = q}} s_\lambda(x) s_\lambda(y).$$

7.24 Enumeration under Group Action

The theory of enumeration under group action, or *Pólya theory*, is a standard topic within enumerative combinatorics which is usually presented without the use of symmetric functions. However, symmetric function theory does lead to a more natural development and is more convenient for certain extensions of the theory. Pólya theory is centered around a certain generating function $Z_G(x)$ for the cycle types of elements of a subgroup G of $\mathfrak{S}_S$, the symmetric group of all permutations of the finite set S. Actually, there is no need at first for G to be a subgroup, so we make the definition for any *subset* of $\mathfrak{S}_S$. (There are in fact interesting results for certain subsets that aren't subgroups; see for instance Exercise 7.111.)

7.24.1 Definition. Let K be a subset of the symmetric group $\mathfrak{S}_S$. Define the *augmented cycle indicator* $\tilde{Z}_K$ of K to be the symmetric function

$$\tilde{Z}_K = \sum_{w \in K} p_{\rho(w)},$$

where $\rho(w)$ denotes the cycle type of w as in Section 7.7. The *cycle indicator* Z_K of K is defined by

$$Z_K = \frac{1}{\#K} \tilde{Z}_K = \frac{1}{\#K} \sum_{w \in K} p_{\rho(w)}.$$

Thus $\tilde{Z}_K$ or Z_K is just a generating function for elements of K according to their cycle type. Note that if $n = \#S$, then $\tilde{Z}_K$ and Z_K are homogeneous of degree n, i.e., $\tilde{Z}_K, Z_K \in \Lambda^n$. In the traditional exposition of Pólya theory mentioned above, the power sum symmetric function p_i is replaced by an indeterminate t_i, and later one substitutes p_i or a specialization of p_i for t_i. (We have done this ourselves in Example 5.2.10.) This approach represents only a change in viewpoint, since the p_i's are algebraically independent. The cycle indicator, regarded as a polynomial in the indeterminates $t_1, t_2, \ldots$, is then also called the *cycle index polynomial* of K. The main result of Pólya theory expresses Z_G in terms of the monomial symmetric functions (i.e., gives a combinatorial interpretation of the scalar product $\langle Z_G, h_\lambda \rangle$) when G is a subgroup of $\mathfrak{S}_S$.

7.24.2 Example. (a) Let S be the set of vertices of a square, and let G be the dihedral group of all (Euclidean) symmetries of the square, acting on the vertex set S. The identity element has cycle indicator p_1^4. The rotations by 90° or 270° have indicator p_4. The rotation by 180° has indicator p_2^2. The horizontal and vertical reflections also have indicator p_2^2. Finally the two diagonal reflections have indicator $p_1^2 p_2$. Hence

$$Z_G = \tfrac{1}{8}\left(p_1^4 + 2p_1^2 p_2 + 3p_2^2 + 2p_4\right).$$

If instead we let G be the group of rotational symmetries of the square, then we would get

$$Z_G = \tfrac{1}{4}\left(p_1^4 + p_2^2 + 2p_4\right).$$

(b) Let V be a p-element vertex set, and let $S = \binom{V}{2}$. The symmetric group $\mathfrak{S}_V$ acts naturally on S, viz., if $w \in \mathfrak{S}_V$ and $\{s,t\} \in S$, then $w \cdot \{s,t\} = \{w \cdot s, w \cdot t\}$. Thus we have a subgroup $\mathfrak{S}_p \cong G \subset \mathfrak{S}_S \cong \mathfrak{S}_{\binom{p}{2}}$. For instance, when $p = 4$ we have

$$Z_G = \tfrac{1}{24}\left(p_1^6 + 9p_1^2 p_2^2 + 8p_3^2 + 6p_2 p_4\right).$$

(c) Let G be the group $\mathfrak{S}_S$ of *all* permutations of the n-element set S, so $G \cong \mathfrak{S}_n$. Let $\lambda \vdash n$. Recall (equation (7.18)) that $n! z_\lambda^{-1}$ is the number of permutations $w \in \mathfrak{S}_n$ of cycle type λ. Hence from equation (7.22) we get $Z_G = h_n$.

Now let $X = \{c_1, c_2, \ldots\}$ be a set of "colors," and let X^S denote the set of all functions $f : S \to X$. Think of f as a "coloring" of the set S, where the element $s \in S$ receives the color $f(s)$. Define the *weight* x^f of $f \in X^S$ by

$$x^f = \prod_{i \geq 1} x_i^{\#f^{-1}(c_i)}.$$

Thus x^f is a monomial of degree $n = \#S$ in the variables $x_1, x_2, \ldots$, which tells us for each i how many elements of S are colored c_i. There is a natural action of G on X^S, viz., if $w \in G$, $f \in X^S$, and $s \in S$, then

$$(w \cdot f)(s) = f(w \cdot s).$$

Let X^S/G denote the set of orbits of this action. In other words, define an equivalence relation $\sim$ on X^S by $f \sim g$ if there exists $w \in G$ such that $g = w \cdot f$. Then the elements of X^S/G are the equivalence classes with respect to $\sim$. Note that if $f \sim g$ then $x^f = x^g$. Hence if $\mathcal{O} \in X^S/G$, then we can define $x^{\mathcal{O}}$ to be x^f for any $f \in \mathcal{O}$. Equivalence classes $\mathcal{O}$ are called *patterns*. The *pattern inventory* (also known by various other names, such as *store enumerator* and *configuration counting series*) of G is the generating function

$$F_G(x) = \sum_{\mathcal{O} \in X^S/G} x^{\mathcal{O}}.$$

Thus the coefficient of a monomial x^α in $F_G(x)$ is the number of orbits $\mathcal{O} \in X^S/G$ of weight x^α. Since the elements of X are all "treated equally," it follows that $F_G(x)$ is a symmetric function. In fact, $F_G \in \Lambda^n$, since by definition F_G is homogeneous of degree n.

7.24.3 Example. (a) Traditionally the elements of S have some combinatorial or geometric structure. For instance, let S be the set of vertices of a square and G the group of dihedral symmetries as in Example 7.24.2(a). Two colorings of the vertices are equivalent if there is a symmetry of the square taking one coloring to the other. Let $\lambda \vdash 4$. The coefficient of m_λ in F_G is the number of inequivalent vertex colorings (or colorings "up to symmetry") using λ_i i's. Here are representatives (orbit members) of each of these inequivalent colorings (using the colors $1, 2, \ldots$):

$$\begin{array}{ccccccccc} 11 & 11 & 11 & 21 & 11 & 12 & 12 & 12 & 13 \\ 11 & 12 & 22 & 12 & 23 & 31 & 34 & 43 & 42 \end{array}$$

Hence

$$F_G = m_4 + m_{31} + 2m_{22} + 2m_{211} + 3m_{1111}.$$

If the group G were instead the group of rotational symmetries of the square, then we would get the additional inequivalent colorings

$$\begin{array}{cccc} 11 & 13 & 14 & 14 \\ 32 & 24 & 23 & 32 \end{array}.$$

Hence in this case

$$F_G = m_4 + m_{31} + 2m_{22} + 3m_{211} + 6m_{1111}.$$

(b) Let V, S, and G be as in Example 7.24.2(b), so $S = \binom{V}{2}$. Let $X = \mathbb{P}$. A function $f \in X^S$ may be regarded as a graph on the vertex set V, allowing multiple edges but not loops. Namely, if $f(\{s,t\}) = j$, then place $j-1$ (indistinguishable) edges between vertices s and t. Two graphs $f, g \in X^S$ are equivalent if and only if there is a permutation w of their vertex set such that w preserves edges, i.e., there are j edges between s and t if and only if there are j edges between $w \cdot s$ and $w \cdot t$. In other words, $f \sim g$ if and only if f and g are *isomorphic* (as graphs). Thus for $\alpha \in \mathrm{Comp}(\binom{p}{2})$ the coefficient of the monomial x^α in F_G is equal to the number of nonisomorphic loopless graphs with p vertices and α_j edges of multiplicity $j-1$. For the case $p = 4$ we have

$$F_G = m_6 + m_{51} + 2m_{42} + 3m_{33} + 2m_{411} + 4m_{321} + 6m_{222}$$
$$+ 5m_{3111} + 9m_{2211} + 15m_{21111} + 30m_{111111}.$$

Moreover, $F_G(1^m)$ is equal to the total number of nonisomorphic loopless graphs with p vertices and all edges of multiplicity at most $m-1$. In particular, $F_G(1,1)$ is the number of nonisomorphic simple graphs (no loops or multiple edges) on p vertices. Note that the specialization $F_G(1^m)$ is obtained by expanding F_G in terms of the power sums p_i and setting each $p_i = m$. More generally, the coefficient of q^j in the polynomial $F_G(1, q, \ldots, q^{m-1})$ is the number of nonisomorphic loopless graphs with p vertices, all edges of multiplicity at most $m-1$, and exactly j edges. Thus for instance when $p = 4$ we get

$$F_G(1,q) = 1 + q + 2q^2 + 3q^3 + 2q^4 + q^5 + q^6,$$

the generating function for nonisomorphic 4-vertex simple graphs by number of edges.

NOTE. Since traditionally the variables $t_1, t_2, \ldots$ of Z_G and F_G correspond to the power sums $p_1, p_2, \ldots$ instead of the *arguments* $x_1, x_2, \ldots$ of the p_i's, what we write as $F_G(1, q, \ldots, q^{m-1})$ is traditionally written

$$F_G\big(1 + q + \cdots + q^{m-1}, 1 + q^2 + \cdots + q^{2(m-1)}, 1 + q^3 + \cdots + q^{3(m-1)}, \ldots\big).$$

(c) Let G be the group $\mathfrak{S}_S$ of *all* permutations of the n-element set S, as in Example 7.24.2(c). For simplicity take $S = [n]$, so $\mathfrak{S}_S = \mathfrak{S}_n$. Two colorings $f, g \in X^S$ are equivalent if and only if their coimages have the same type, i.e., the multisets $\{\#f^{-1}(c) : c \in X\}$ and $\{\#g^{-1}(c) : c \in X\}$ are the same. Hence the coefficient of every monomial x^α of degree n in $F_{\mathfrak{S}_n}$ is equal to 1, so

$$F_{\mathfrak{S}_n} = \sum_{\lambda \vdash n} m_\lambda = h_n.$$

We have defined two symmetric functions Z_G and F_G associated with the permutation group G. The cycle indicator Z_G is defined in terms of the power sum symmetric functions p_λ, while the pattern inventory is defined in terms of the

monomial symmetric functions m_λ. The main result of Pólya theory is that these two symmetric functions are equal.

7.24.4 Theorem. *Let S be a finite set. For any subgroup G of $\mathfrak{S}_S$ we have $Z_G = F_G$.*

The proof is based on a simple but fundamental result on permutation groups known as *Burnside's lemma* (though see the Notes for the correct attribution).

7.24.5 Lemma. *Let Y be a finite set and G a subgroup of $\mathfrak{S}_Y$. For each $w \in G$ let*

$$\mathrm{Fix}(w) = \{y \in Y : w(y) = y\},$$

so $\#\mathrm{Fix}(w)$ is the number of cycles of length one in the permutation w. Let Y/G be the set of orbits of G. Then

$$\#(Y/G) = \frac{1}{\#G}\sum_{w\in G}\#\mathrm{Fix}(w).$$

In other words, the average number of elements of Y fixed by an element of G is equal to the number of orbits.

Proof. For $y \in Y$ let $G_y = \{w \in G : w(y) = y\}$, the *stabilizer* of y. Then

$$\begin{aligned}\frac{1}{\#G}\sum_{w\in G}\mathrm{Fix}(w) &= \frac{1}{\#G}\sum_{w\in G}\sum_{\substack{y\in Y\\ w(y)=y}} 1\\ &= \frac{1}{\#G}\sum_{y\in Y}\sum_{\substack{w\in G\\ w(y)=y}} 1\\ &= \frac{1}{\#G}\sum_{y\in Y}\#G_y.\end{aligned}$$

Let $Gy = \{w(y) : w \in G\}$, the orbit of G containing y. The multiset of elements $w(y)$, $w \in G$, contains every element in the orbit Gy the same number of times [why?], viz., $\#G/\#Gy$ times. Thus y occurs $\#G/\#Gy$ times among the elements $w(y)$, so

$$\frac{\#G}{\#Gy} = \#G_y.$$

Hence

$$\frac{1}{\#G}\sum_{w\in G}\#\mathrm{Fix}(w) = \frac{1}{\#G}\sum_{y\in Y}\frac{\#G}{\#Gy}$$
$$= \sum_{y\in Y}\frac{1}{\#Gy}.$$

For a fixed orbit $\mathcal{O} \in Y/G$, we have $Gy = \mathcal{O}$ if and only if $y \in \mathcal{O}$. Hence the term $1/\#\mathcal{O}$ appears $\#\mathcal{O}$ times in the last sum above, so the sum is equal to the number of orbits. □

Proof of Theorem 7.24.4. Let $\alpha = (\alpha_1, \alpha_2, \ldots)$ be a weak composition of n, and let $\mathcal{C}_\alpha$ denote the set of all "colorings" $f \in X^S$ with color c_j used α_j times. The set $\mathcal{C}_\alpha$ is invariant under the action of G on X^S. Let w_α denote the action of w on $\mathcal{C}_\alpha$. We want to apply Burnside's lemma (Lemma 7.24.5) to compute the number of orbits, so we need to find $\#\mathrm{Fix}(w_\alpha)$.

In order for $f \in \mathrm{Fix}(w_\alpha)$, we must color S so that (a) in any cycle of w, all the elements get the same color, and (b) the color c_j appears α_j times. It follows that

$$\#\mathrm{Fix}(w_\alpha) = [x^\alpha]\prod_j p_j^{m_j(w)} = [x^\alpha]p_{\rho(w)},$$

where $m_j(w)$ is the number of cycles of w of length j. Hence

$$p_{\rho(w)}(x) = \sum_\alpha \#\mathrm{Fix}(w_\alpha)x^\alpha.$$

Now sum over all $w \in G$ and divide by $\#G$. The left-hand side becomes Z_G, while by Burnside's lemma the right-hand side becomes F_G. □

When we put $x = 1^m$ in Theorem 7.24.4 we get the following result (which can also be obtained directly from Burnside's theorem).

7.24.6 Corollary. *Let $N_G(m)$ be the total number of inequivalent colorings of S from a set of m colors. Then*

$$N_G(m) = \frac{1}{\#G}\sum_{w\in G} m^{c(w)},$$

where $c(w)$ is the number of cycles of w.

NOTE (for algebraists). Let G be a subgroup of $\mathfrak{S}_S$ as above. Let $r = [\mathfrak{S}_S : G]$, the index of G in $\mathfrak{S}_S$. The group $\mathfrak{S}_S$ acts on the (right) cosets of G, defining a (transitive) permutation representation of $\mathfrak{S}_S$. Representing a permutation by the corresponding permutation matrix gives a linear representation $\sigma^G : \mathfrak{S}_S \to \mathrm{GL}(r, \mathbb{C})$. This linear representation is equivalent to the induced representation

$1_G^{\mathfrak{S}_S}$ (the induction of the trivial representation of G to $\mathfrak{S}_S$). Let $\chi^G : \mathfrak{S}_S \to \mathbb{C}$ denote the character of this representation, i.e., $\chi^G(w) = \operatorname{tr} \sigma^G(w)$. If $w \in \mathfrak{S}_S$ and $\rho(w) = \lambda$ (the cycle type of w), then

$$\chi^G(w) = \frac{\#\mathfrak{S}_S}{\#\mathfrak{S}^\lambda} \cdot \frac{\#(\mathfrak{S}^\lambda \cap G)}{\#G} = \frac{z_\lambda \cdot \#(\mathfrak{S}^\lambda \cap G)}{\#G}, \tag{7.118}$$

where $\mathfrak{S}^\lambda$ denotes the subset (conjugacy class) of $\mathfrak{S}_S$ of all permutations of cycle type λ. It follows from (7.118) that

$$\operatorname{ch}(\chi^G) = Z_G, \tag{7.119}$$

the cycle indicator of G, where ch is the characteristic map of Section 7.18. Hence Pólya theory is closely related to the interaction between the representation theory of the symmetric group and the theory of symmetric functions.

Write

$$\chi^G = \sum_{\lambda \vdash n} a_\lambda \chi^\lambda$$

as the decomposition of the character χ^G in terms of irreducible characters χ^λ. Hence a_λ is the multiplicity of χ^λ in χ^G, and so $a_\lambda \in \mathbb{N}$. Since the map ch is linear and $\operatorname{ch}(\chi^\lambda) = s_\lambda$, there follows from equation (7.119) that

$$Z_G = \sum_{\lambda \vdash n} a_\lambda s_\lambda. \tag{7.120}$$

Thus we have an algebraic interpretation of the expansion of Z_G in terms of Schur functions, showing in particular that the coefficients $a_\lambda = \langle Z_G, s_\lambda \rangle$ are nonnegative integers, or equivalently that Z_G is s-integral (which is trivial) and s-positive. No purely combinatorial or "formal" proof of the s-positivity of Z_G is known; all known proofs (which are essentially equivalent) use representation theory. (The fact that a_λ is an integer is immediate from Theorem 7.24.4.) There are a myriad of other "positivity theorems" in the theory of symmetric functions whose only known proofs use representation theory.

Notes

A good source of information for the early history of symmetric functions, such as the fundamental theorem of symmetric functions (Theorem 7.4.4) and the symmetry of the matrix $(M_{\lambda\mu})$ (Corollary 7.4.2), is the article **[154]** by Karl Theodor Vahlen. In particular, the first published work on symmetric functions is due to Albert Girard **[49]** in 1629, who gives an explicit formula expressing p_n in terms of the e_λ's (which we can obtain by equating coefficients of t^n in the formula $\sum_{n\geq 1} \frac{1}{n}(-1)^{n-1} p_n t^n = \log \sum_{k\geq 0} e_k t^k$). Other early researchers on symmetric

functions include Gabriel Cramer, Francesco Faà di Bruno, Isaac Newton, and Edward Waring.

The earliest reference to Schur functions is the 1815 paper [**18**] of Augustin Louis Cauchy. (Cauchy's paper was submitted for publication in 1812.) He defines Schur functions (though of course not by that name) in the variables $x_1, \ldots, x_n$ as the bialternants of Theorem 7.15.1 and proves that they are indeed symmetric polynomials, and that (using our notation) $s_1(x_1, \ldots, x_n) = x_1 + \cdots + x_n$ and (more trivially) $s_{1^n}(x_1, \ldots, x_n) = x_1 \cdots x_n$. The next paper of interest to us is that of Karl Gustav Jacob Jacobi [**64**] in 1841, in which he states without proof the Jacobi–Trudi identity (Theorem 7.16.1) in the case $\mu = \emptyset$, i.e., for ordinary Schur functions s_λ (defined as bialternants). In 1864 Jacobi's student Nicolò Trudi [**153**] gave a complete proof of the Jacobi–Trudi identity. The dual Jacobi–Trudi identity (Corollary 7.16.2) is due to Hans Eduard von Nägelsbach [**108**] in 1871, and was given a simpler proof in 1875 by Carl Franz Albert Kostka [**73**]. Our first proof of the Jacobi–Trudi formula (based on the theory of nonintersecting lattice paths) follows Ira Martin Gessel and Gérard Xavier Viennot [**47**][**2.5**]. See [**151**, Ch. 4.5] for an exposition. Our second proof comes from Ian G. Macdonald [**96**, Ch. I, (5.4)].

The expansion of $\prod(1 - x_i y_j)^{-1}$ in terms of Schur functions (Theorem 7.12.1) is universally attributed to Cauchy and is therefore called the "Cauchy identity." We have been unable, however, to find a clear statement of this identity in the work of Cauchy. On the other hand, the Cauchy identity is an almost trivial consequence of two results of Cauchy. The first result is the Binet–Cauchy formula for the determinant of the product of an $m \times n$ matrix and an $n \times m$ matrix. (For an interesting discussion of Cauchy's precise contribution to this formula, see [**106**, pp. 92–131]. It is at any rate clear from Cauchy's later work that he was adept at its use.) The second result of Cauchy [**19**, eqn. (10)] is the determinant evaluation (stated slightly differently)

$$\det\left(\frac{1}{1 - x_i y_j}\right)_{i,j=1}^{n} = \frac{a_\delta(x) a_\delta(y)}{\prod(1 - x_i y_j)}, \tag{7.121}$$

where a_δ is the Vandermonde determinant (equation (7.55)). Applying the Binet–Cauchy formula to the product $A(x)A(y)^t$, where

$$A(z) = \begin{bmatrix} 1 & z_1 & z_1^2 & \cdots \\ 1 & z_2 & z_2^2 & \cdots \\ \vdots & \vdots & \vdots & \\ 1 & z_n & z_n^2 & \cdots \end{bmatrix},$$

gives

$$\det\left(\frac{1}{1 - x_i y_j}\right)_{i,j=1}^{n} = \sum_{\ell(\lambda) \le n} a_{\lambda+\delta}(x) a_{\lambda+\delta}(y).$$

Hence from (7.121) we have

$$\frac{1}{\prod_{i,j=1}^{n}(1-x_iy_j)} = \sum_{\ell(\lambda)\le n} \frac{a_{\lambda+\delta}(x)}{a_\delta(x)} \cdot \frac{a_{\lambda+\delta}(y)}{a_\delta(y)},$$

and Cauchy's identity follows from Theorem 7.15.1 (which was Cauchy's definition of Schur functions).

Kostka [**74**][**75**] was the first person to consider the expansion of Schur functions into monomials, whence the term "Kostka number" for the numbers $K_{\lambda\mu}$ in the expansion $s_\lambda = \sum_\mu K_{\lambda\mu} m_\mu$ (equation (7.35)). (Foulkes [**36**, p. 85] suggested calling the matrix $(K_{\lambda\mu})$ the *Kostka matrix*.) In [**74**] Kostka gives a table of the Kostka numbers $K_{\lambda\mu}$ and the entries $(K^{-1})_{\lambda\mu}$ of the inverse Kostka matrix for $\lambda, \mu \vdash n \le 8$. In [**75**] he extends the tables up to $n = 11$. Kostka asserts that the numbers in his tables also give the expansion of the h_μ's and m_μ's in terms of the s_λ's; this assertion is equivalent to the orthonormality of the s_λ's (Corollary 7.12.2).

Oscar Howard Mitchell [**103**] looked further at Kostka numbers in 1882. He showed that they were nonnegative without obtaining an explicit combinatorial interpretation of them, and he evaluated $s_\lambda(1^n)$ in the form obtained by letting $q \to 1$ in equation (7.105). A simpler proof of this evaluation of $s_\lambda(1^n)$ was later given by William Woolsey Johnson [**66**, §13]. Some efforts were subsequently made to find combinatorial interpretations of Kostka numbers in special cases, a typical example being Thomas Muir [**105**]. The first explicit statement of which we are aware that the Kostka number $K_{\lambda\mu}$ counts SSYTs of shape λ and type μ is due to Dudley Ernest Littlewood [**85**, Thm. VI] (see also [**88**, Ch. 10.1, IX]). The definition of the Gelfand–Tsetlin patterns of equation (7.37) was given by Israel M. Gelfand and Mikhail L. Tsetlin [**45**, (3)] in connection with the representation theory of the Lie algebra $\mathfrak{sl}(n, \mathbb{C})$.

Skew Schur functions were first investigated by Nägelsbach [**108**] and Alexander Craig Aitken [**1**][**2**], in the form given by equation (7.68). Aitken proved what in our notation is the formula $\omega s_{\lambda/\nu} = s_{\lambda'/\nu'}$ (Theorem 7.15.6), and in a later paper [**3**] gave the combinatorial interpretation of $s_{\lambda/\mu}$ in terms of skew SSYTs. The connection between skew Schur functions and Littlewood–Richardson coefficients (Theorem 7.15.4) appears in [**88**, eqn. VIII, p. 110].

While the work described above was being carried out, a completely independent but ultimately equivalent avenue of research was being developed by a group of geometers. Certain enumerative questions involving intersections of subspaces of a vector space were reduced to algebraic computations formally the same as basic results in the theory of Schur functions. This general approach to enumerative geometry was first developed by Hermann Cäsar Hannibal Schubert (see [**68**] for references) and is now known as the *Schubert calculus*. It is not our intention to explain the Schubert calculus and its connections with symmetric functions here, but we will briefly mention the main highlights. The geometric result equivalent to Pieri's rule (Theorem 7.15.7) was given in 1893 by Mario Pieri [**118**]. A

determinantal formula expressing a Schubert cycle in terms of special Schubert cycles was published in 1903 by Giovanni Zeno Giambelli [**48**]. This formula was formally the same as the Jacobi-Trudi identity (Theorem 7.16.1), thereby establishing the formal equivalence (though no one realized it yet) of the Schubert calculus with the algebra of Schur functions. The work of classical geometers such as Schubert, Pieri, and Giambelli on the Schubert calculus was vindicated rigorously by Charles Ehresmann [**28**], Bartel Leendert van der Waerden [**158**], William Vallance Douglas Hodge [**61**][**62**] *et al.* The formal equivalence between Schubert calculus and the algebra of Schur functions was first pointed out by Léonce Lesieur [**84**] in 1947. Conceptual explanations of this seeming "coincidence" were first given by Geoffrey Horrocks [**63**] and James B. Carrell [**17**]. More details of this history are discussed by William Fulton [**41**, pp. 278–279]. For three surveys of the Schubert calculus, the third one focusing on the connections with combinatorics, see [**68**][**69**][**148**].

The idea of unifying much of the theory of symmetric functions using linear algebra (scalar product, dual bases, involution, etc.) is due to Philip Hall in his important paper [**54**]. This paper was overlooked (undoubtedly because of its obscure place of publication) until an exposition (with most of the missing proofs filled in) was given by Stanley [**145**]. Stanley learned of Hall's paper from Robert James McEliece, who studied with Hall during the 1964–65 academic year in Cambridge (England). A later exposition of Hall's work was given by Macdonald [**94**].

The RSK algorithm (known by a variety of other names: either "correspondence" or "algorithm" in connection with some subset of the names Robinson, Schensted, and Knuth) was first described, in a rather vague form, by Gilbert de Beauregard Robinson [**131**, §5], as a tool in an attempted proof of the Littlewood–Richardson rule (Appendix 1, §A1.3). (See the Notes to Appendix 1 for the history of the Littlewood–Richardson rule.) The RSK algorithm was later rediscovered by Craige Eugene Schensted (see below), but no one actually analyzed Robinson's work until this was done by Marc A. A. van Leeuwen [**82**, §7]. It is interesting to note that Robinson says in a footnote on page 754 that "I am indebted for this association I to Mr. D. E. Littlewood." Van Leeuwen's analysis makes it clear that "association I" gives the recording tableau Q of the RSK algorithm $w \xrightarrow{\text{RSK}} (P, Q)$. Thus it might be correct to say that if $w \in \mathfrak{S}_n$ and $w \xrightarrow{\text{RSK}} (P, Q)$, then the definition of P is due to Robinson, while the definition of Q is due to Littlewood.

No further work related to Robinson's construction was done until Schensted published his seminal paper [**136**] in 1961. (For some information about the unusual life of Schensted, see [**7**].) Schensted's purpose was the enumeration of permutations in $\mathfrak{S}_n$ according to the length of their longest increasing and decreasing subsequences. For further information see the discussion of Section 7.23 below. According to Knuth [**71**, p. 726], the connection between the work of Robinson and that of Schensted was first pointed out by Marcel Paul Schützenberger, though as mentioned above the first person to describe this connection precisely was van Leeuwen.

Robinson states on page 755 of his paper **[131]** on the RSK algorithm that "it is not difficult to see" that if $w \xrightarrow{\text{RSK}} (P, Q)$, then $w^{-1} \xrightarrow{\text{RSK}} (Q, P)$. No indication of a proof of this fundamental result (our Theorem 7.13.1) is given. A proof was finally given by Schützenberger **[140]** in 1963. Schützenberger was the first to realize the great significance of Schensted's work for the theory of symmetric functions and the symmetric group. Our first proof of Theorem 7.13.1 follows Donald Ervin Knuth [**71**, §4][**72**, Ch. 5.1.4]. Corollary 7.13.8, which we derived from Theorem 7.13.1, was first proved by Issai Schur **[138]** (repeated in [**5.53**, VII.47]).

The theory of growth diagrams, which we used for our second proof of Theorem 7.13.1, was developed by Sergey V. Fomin **[31][32][33][34]**. Some further work was done by Thomas W. Roby **[132][133]**. Before Fomin a different "geometric" theory of the RSK algorithm had been developed by Viennot **[155][156]**.

The extension of the RSK algorithm from permutations to arbitrary sequences of nonnegative integers (or from permutation matrices to $\mathbb{N}$-matrices of finite support) is due to Knuth **[71]**. Although, as pointed out by Lemma 7.11.6, this extension is actually equivalent to the original case, it is essential to use the more general form when dealing with symmetric functions. Thus for instance we obtained a direct bijective proof of the Cauchy identity (Theorem 7.12.1), as first done by Knuth [**71**, p. 726]. Knuth's paper deals with a number of further topics related to the RSK algorithm, in particular, a proof of the symmetry result Theorem 7.13.1 (working directly with $\mathbb{N}$-matrices of finite support and not reducing to the case of permutations), the definition and basic properties of the dual RSK algorithm of Section 7.14, and the definition and basic properties of Knuth equivalence, as discussed in Appendix 1. Some further variations of the RSK algorithm were given by William H. Burge **[13]** (see Exercises 7.28(c,e) and 7.29(a,b)). For good overviews of more recent work related to the RSK algorithm, see **[134][83]**.

By this time the work described above had entered the general consciousness of algebraic and enumerative combinatorics, and the floodgates were opened. We will not attempt a survey of the enormous amount of more recent work done on symmetric functions, Young tableaux, the RSK algorithm, etc., but we will give some references to work discussed in the text. For further developments, see the end of these Notes, the Exercises to this chapter, and the Notes to Appendix 1.

Standard Young tableaux (SYT) were first enumerated by Percy Alexander MacMahon [**99**, p. 175] (see also [**101**, §103]). MacMahon formulated his result in terms of the ballot sequences or lattice permutations of Proposition 7.10.3(c,d), and stated the result not in terms of the product of hook lengths as in Corollary 7.21.6, but rather using the right-hand side of the case $q = 1$ of equation (7.101). The formulation in terms of hook lengths is due to James Sutherland Frame and appears first in the paper [**38**, Thm. 1] of Frame, Robinson, and Robert McDowell Thrall; hence it is sometimes called the "Frame–Robinson–Thrall hook-length formula." (The actual definition of standard Young tableaux is due to Alfred Young [**162**, p. 258].) Independently of MacMahon, Ferdinand Georg Frobenius [**5.27**, eqn. (6)] obtained the same formula for the degree of the irreducible character χ^λ

of $\mathfrak{S}_n$ as MacMahon obtained for the number of lattice permutations of type λ. Frobenius was apparently unaware of the combinatorial significance of deg χ^λ, but Young showed in [**162**, pp. 260–261] that deg χ^λ was the number of SYT of shape λ, thereby giving an independent proof of MacMahon's result. (Young also provided his own proof of MacMahon's result in [**162**, Thm. II].) A number of other proofs of the hook-length formula were subsequently found. Curtis Greene, Albert Nijenhuis, and Herbert Saul Wilf [**51**] gave an elegant probabilistic proof. The proof we gave in Section 7.22 based on the Hillman–Grassl correspondence appears in [**60**] and shows very clearly the role of hook lengths, though the proof is not completely bijective. A bijective version was later given by Christian Krattenthaler [**76**]. Completely bijective proofs of the hook-length formula were first given by Deborah Franzblau and Doron Zeilberger [**39**] and by Jeffrey Brian Remmel [**127**]. An exceptionally elegant bijective proof was later found by Jean-Christophe Novelli, Igor Pak, and Alexander V. Stoyanovskii [**113**].

For more information on the Hopf algebra approach to symmetric functions mentioned at the end of Section 7.15 see [**163**].

The determinantal formula (7.71) for $f^{\lambda/\mu}$ is due to Aitken [**3**, p. 310], who deduced it just as we have done from the Jacobi–Trudi identity for $s_{\lambda/\mu}$ (Theorem 7.16.1). Aitken's result was rediscovered by Walter Feit [**30**]. A generalization due to Germain Kreweras is given by Exercise 3.63.

The theory of representations of finite groups was developed by Frobenius; see [**24**][**56**][**57**][**58**][**59**] for an interesting discussion of this development. In particular, Frobenius computed the irreducible characters of $\mathfrak{S}_n$ (in the form given by Corollary 7.17.4 or equation (7.77)) in [**5.27**]. Much subsequent work on the representation theory of $\mathfrak{S}_n$ was done by Alfred Young; see Sagan [**135**] for a nice exposition of Young's work and its connection with symmetric functions. The Murnaghan–Nakayama rule (regarded as the formula (7.75) for $\chi^\lambda(\mu)$) is actually due to Littlewood and Archibald Read Richardson [**89**, §11]. Statements of this rule by Francis Dominic Murnaghan and Tadasi Nakayama are given in [**107**, (13)] and [**109**, §9].

The definition of quasisymmetric functions is due to Gessel, though they had appeared implicitly in earlier work. Gessel used quasisymmetric functions to prove such results as our Corollaries 7.23.6 and 7.23.8. The basic results on (P, ω)-partitions used here (Theorem 7.19.4 and Corollary 7.19.5) were given by Stanley [**3.28**, Ch. 2][**3.29**, Ch. 1], though not using the language of quasisymmetric functions. Proposition 7.19.11 seems to be due to George Lusztig (unpublished) and Stanley [**149**, Prop. 4.11]. Further work on quasisymmetric functions appears e.g. in [**26**][**102**] and the references given there.

Plane partitions were discovered by MacMahon in a series of papers which were not appreciated until much later. (See MacMahon's book [**101**, §§IX and X] for an exposition of his results.) MacMahon's first paper dealing with plane partitions was [**98**]. In Article 43 of this paper he gives the definition of a plane partition (though not yet with that name), and then goes on to discuss the six aspects of a plane partition. In Article 51 he conjectures that the generating function for plane

partitions is the product

$$(1-x)^{-1}(1-x^2)^{-2}(1-x^3)^{-3}(1-x^4)^{-4}\cdots$$

(our Corollary 7.20.3). He also suggests (but doesn't call it a conjecture) that the generating function for r-dimensional partitions (whose diagram would be a finite order ideal of $\mathbb{N}^{r+1}$) is

$$\prod_{i\geq 1}\left(1-x^i\right)^{-\binom{i+r-2}{r-1}}. \tag{7.122}$$

MacMahon apparently never realized that this generating function, even for $r=3$, is incorrect, though he does mention in [**101**, vol. 2, footnote on p. 175] that an even stronger result is false. (When $r=3$, the smallest exponent n for which the coefficient of x^n in (7.122) fails to be the number of 3-dimensional partitions of n is $n=6$.) The incorrectness of (7.122) was first shown by Atkin *et al.* [**6**] and later by E. M. Wright [**160**]. V. S. Nanda [**110**][**111**] erroneously assumes (7.122) to be correct for $r=3$, stating in [**110**, p. 593] that

MacMahon has not given a rigorous derivation of the generating function for solid partitions. But a simple reasoning as in the case of plane partitions leads to the generating function

$$\frac{1}{(1)(2)^3(3)^6\cdots(s)^{\frac{1}{2}(s^2+s)}\cdots}$$

for solid partitions when there is no restriction on part magnitude.

Further computations by Knuth [**70**] show how useless it seems to write the generating function for 3-dimensional partitions in the form $\prod_{i\geq 1}(1-x^i)^{a_i}$. Returning to the paper [**98**] of MacMahon, in Article 52 he conjectures our Theorem 7.20.1, Corollary 7.20.2, and finally Theorem 7.21.7 (which includes all the previous conjectures as special cases). MacMahon goes on in Articles 56–62 to prove his conjecture in the case of plane partitions with at most 2 rows and c columns (the case $r=2$ of our Theorem 7.20.1), mentioning on page 662 that an independent solution was obtained by Andrew Russell Forsyth. (Though a publication reference is given to Forsyth's paper, apparently it never actually appeared.)

We will not attempt to describe MacMahon's subsequent work on plane partitions, except to say that the culmination of his work appears in [**101**, Art. 495], in which he proves his main conjecture from his first paper [**98**] on plane partitions, viz., our Theorem 7.21.7. MacMahon's proof is quite lengthy and indirect. We can regard a plane partition whose shape is contained in the partition $\lambda \vdash p$ as a P_λ-partition (in the sense of Section 4.5), where P_λ is the poset defined after Corollary 7.19.5. (We regard P_λ as a natural partial order on $[p]$, as in Section 4.5.) MacMahon anticipates the theory of P-partitions (as pointed out in the Notes to Chapter 4) by essentially establishing Exercise 4.24(b) for $P=P_\lambda$. He manages to convert this expression into a determinant, and then to evaluate the determinant when $\lambda=\langle c^r\rangle$.

Some interesting recent work on the shape of the diagram of a "typical" plane partition fitting in an $r \times c \times t$ box was done by Henry Cohn, Michael Larsen, and James Propp [**23**].

The theory of plane partitions stayed rather dormant until the early 1960s, when Basil Gordon and his students made a number of new contributions (but without using symmetric functions). For further discussion and references to the work of Gordon, see [**146**]. Also about this time Carlitz [**16**] gave a simpler proof of MacMahon's main result (our Theorem 7.21.7). (Limiting cases had earlier been given simpler proofs by Chaundy [**21**].) Then in 1972 Edward Anton Bender and Knuth [**8**], in an important paper, showed the connection between the theory of symmetric functions and the enumeration of plane partitions. They gave simple proofs, based on the RSK algorithm, of many of the results (and some generalizations) of Gordon *et al.*, as well as the first bijective proof (the same proof that we give) of our Theorem 7.20.1 in the case $q = 1$. The introduction of the variable q in Theorem 7.20.1 and related results to keep track of the trace of a plane partition is due to Stanley [**146**, Thm. 19.3][**147**].

The "hook-content formula" for $s_\lambda(1, q, \ldots, q^{n-1})$ (Theorem 7.21.2) was first stated explicitly by Stanley [**3.28**, Thm. V.2.3][**3.29**, Prop. 21.3][**146**, Thm. 15.3]. Earlier a less explicit statement (using the right-hand side of our equation (7.101) instead of the left-hand side) was given by Littlewood and Richardson [**90**, Thm. I] [**88**, I. on p. 124]. A bijective proof based on an involution principle argument was given by Remmel and Roger Whitney [**128**]. Krattenthaler [**77**] then gave a bijective proof not involving the involution principle, and generalized it in [**78**]. Finally Krattenthaler [**79**] gave a bijective proof of Theorem 7.21.2 analogous to the Novelli–Pak–Stoyanovskii proof [**113**] of the hook-length formula.

The generating function for symmetric plane partitions with at most r rows (the case $q = 1$ of Theorem 7.20.4) was conjectured by MacMahon [**100**, p. 153][**101**, Art. 520] and first proved by Gordon [**50**]. Bender and Knuth [**8**, pp. 42–43] give the same proof as ours based on the RSK algorithm. MacMahon actually makes a stronger conjecture, viz., an explicit formula for the generating function for symmetric plane partitions with at most r rows and largest part at most m (the symmetric analogue of our Theorem 7.21.7). This result proved to be considerably less tractable than the unrestricted (i.e., nonsymmetric) case; it was first proved by George E. Andrews [**4**][**5**]. A subsequent proof based on the Weyl character formula for type B_n was given by Macdonald [**92**, Exam. I.5.17, p. 52][**96**, Exam. I.5.17, pp. 84–85]. A somewhat different proof based on representation theory is due to Robert Alan Proctor [**121**, Prop. 7.2]. For further information related to the enumeration of symmetry classes of plane partitions, see the solution to Exercise 7.103(b).

The generating function (7.112) for reverse plane partitions of a fixed shape was first obtained by Stanley [**3.28**, Cor. V.2.6, p. 174][**146**, Prop. 18.3]; the proof is the same as our first proof of Theorem 7.22.1 (based on symmetric functions). The elegant bijective proof given after our first proof is due to Abraham P. Hillman and

Richard M. Grassl [**60**]. A different bijective proof was later given by Remmel and Whitney [**129**].

The fundamental result relating the RSK algorithm to descent sets (Lemma 7.23.1) is due to Schützenberger [**140**, Remarque 2] and was independently discovered later by Herbert Owen Foulkes [**37**, Thm. 8.1]. Foulkes anticipated Theorem 7.23.2 and its corollaries, but the first explicit statement of results of this nature is due to Gessel [**46**]. The basic connection between the RSK algorithm and increasing and decreasing subsequences (viz., that if $w \xrightarrow{\text{RSK}} (P, Q)$ where P and Q have shape λ, then $\text{is}(w) = \lambda_1$ and $\text{ds}(w) = \lambda'_1$) is the main result of Schensted [**136**]. Schensted's purpose in writing his paper was to obtain a formula for the number of $w \in \mathfrak{S}_n$ satisfying $\text{is}(w) = p$ and $\text{ds}(w) = q$, which he gave as his Theorem 7.3 (our Corollary 7.23.18). Theorems 7.23.13 and 7.23.17 are due to Greene; see Appendix 1 for further details. Example 7.23.19(a) is a famous result of Paul Erdős and George Szekeres [**29**, eqn. (8)] which was later given an elegant simple proof by Abraham Seidenberg [**141**]. Example 7.23.19(b) was posed as a problem by Stanley Rabinowitz [**122**] and solved by Stanley [**144**].

Burnside's lemma (Lemma 7.24.5) was actually first stated and proved by Frobenius [**40**, end of §4]. Frobenius in turn credits Cauchy [**20**, p. 286] for proving the lemma in the transitive case. Burnside, in the first edition of his book [**14**, §§118–119], attributes the lemma to Frobenius, but in the second edition [**15**] this citation is absent. For more on the history of Burnside's lemma, see [**112**] and [**161**]. Many authors (e.g., [**67**]) now call this result the Cauchy–Frobenius lemma. The cycle indicator $Z_G(x)$ (where G is a subgroup of $\mathfrak{S}_n$) was first considered by J. Howard Redfield [**124**], who called it the *group reduction function*, denoted $\text{Grf}(G)$. George Pólya [**119**] independently defined the cycle indicator, proved the fundamental Theorem 7.24.4, and gave numerous applications. For an English translation of Pólya's paper, see [**120**]. Much of Pólya's work was anticipated by Redfield. For interesting historical information about the work of Redfield and its relation to Pólya theory, see [**53**][**55**][**91**][**125**] (all in the same issue of *Journal of Graph Theory*). Subsequent to Pólya's work there have been a huge number of expositions, applications, and generalizations of Pólya theory. We mention here only the nice survey [**12**] by Nicolaas Govert de Bruijn and the paper [**123**] by Ronald C. Read, who was the first to consider the relevance of Schur functions to Pólya theory.

Appendix 1 has its own Notes, so now we discuss Appendix 2. The main result of Appendix 2, the classification of the rational representations of $\text{GL}(n, \mathbb{C})$ and the determination of their characters (Theorem A2.4), appears in the masterful doctoral dissertation of Schur [**137**]. He later gave a simpler proof [**139**] along the lines of the proof we have sketched. A slight refinement was given by Hermann Weyl [**159**, Thm. 4.4.C]. For a modern treatment of the work of Schur, see for instance [**96**, Ch. I, App. A]. Plethysm was introduced by D. E. Littlewood [**86**, p. 329] (see also [**88**, p. 289]). The term "plethysm" was suggested to Littlewood [**87**, p. 274] by M. L. Clark, after the Greek word *plethysmos* ($\pi\lambda\eta\theta\upsilon\sigma\mu\acute{o}\varsigma$) for "multiplication."

The connection between plethysm and wreath products (Theorem A2.8) is implicit in the determination of the characters of the wreath product $G\S\mathfrak{S}_n$ by Wilhelm Specht [**143**]. The special case when $\chi = 1_G^{\mathfrak{S}_k}$ and $\theta = 1_H^{\mathfrak{S}_n}$ for subgroups G and H of $\mathfrak{S}_k$ and $\mathfrak{S}_n$, respectively, is equivalent to Pólya's derivation of the cycle index of his so-called *Kranz group* (*Kranzgruppe*) [**119**, pp. 178–180]. Explicit statements of Theorem A2.8 (though not stated in the language of symmetric functions) may be found in [**93**, Rmk. 6.9][**65**, Ch. 5.4][**96**, App. A, (6.2)].

What should the reader who is interested in learning further about symmetric functions do next? An important topic, not treated at all here, is the myriad generalizations and variations of Schur functions. We present a short list, with some basic references (which are only provided as a means of entry into the subjects), of some of these generalizations.

- Hall–Littlewood symmetric functions [**96**, Ch. III]
- Shifted Schur functions (corresponding to shifted shapes), also called Schur P- and Q-functions [**52**][**96**, Ch. III.8]
- Super-Schur functions [**9**][**96**, Exam. I.3.23, pp. 58–60]
- Zonal symmetric functions [**96**, Ch. VII]
- Jack symmetric functions [**96**, Ch. VI.10][**150**]
- Macdonald symmetric functions [**43**][**96**, Ch. VI][**115**]
- Wreath-product Schur functions $s_{\lambda^{(1)}} \otimes \cdots \otimes s_{\lambda^{(k)}}$ [**96**, Ch. I, App. B][**114**]
- Orthogonal and symplectic Schur functions [**152**]
- Flag Schur functions (or multi-Schur functions) [**157**][**95**, Ch. III][**126**]
- Factorial Schur functions and variations [**10**][**11**][**22**][**97**, §§4–6, 9][**104**]
- Shifted Schur functions (corresponding to "shifted" variables) [**116**][**117**]
- Noncommutative Schur functions of Gelfand, Krob, Lascoux, Leclerc, Retakh, and Thibon [**44**][**80**][**27**][**81**]
- Noncommutative Schur functions of Fomin and Greene [**35**]
- Modular Schur functions (implicit in [**25**, §3.7])
- $\mathrm{GL}(n, \mathbb{F}_q)$-invariant Schur functions of Macdonald [**97**, §7]

References

7.1. A. C. Aitken, On determinants of symmetric functions, *Proc. Edinburgh Math. Soc. (2)* **1** (1929), 55–61.

7.2. A. C. Aitken, Note on dual symmetric functions, *Proc. Edinburgh Math. Soc. (2)* **2** (1931), 164–167.

7.3. A. C. Aitken, The monomial expansion of determinantal symmetric functions, *Proc. Royal Soc. Edinburgh (A)* **61** (1943), 300–310.

7.4. G. E. Andrews, MacMahon's conjecture on symmetric plane partitions, *Proc. Nat. Acad. Sci. U.S.A.* **74** (1977), 426–429.

7.5. G. E. Andrews, Plane partitions (I): The MacMahon conjecture, in *Studies in Foundations and Combinatorics*, Advances in Math. Suppl. Studies, **1**, Academic Press, New York/London, 1978, pp. 131–150.

7.6. A. O. L. Atkin, P. Bratley, I. G. Macdonald, and J. K. S. McKay, Some computations for m-dimensional partitions, *Proc. Cambridge Phil. Soc.* **63** (1967), 1097–1100.

7.7. E. A. Beem, Craige and Irene Schensted don't have a car in the world, *Maine Times*, 12 March, 1982, pp. 20–21.

7.8. E. A. Bender and D. E. Knuth, Enumeration of plane partitions, *J. Combinatorial Theory* **13** (1972), 40–54.

7.9. A. Berele and J. B. Remmel, Hook flag characters and their combinatorics, *J. Pure Appl. Algebra* **35** (1985), 225–245.

7.10. L. C. Biedenharn and J. D. Louck, A new class of symmetric polynomials defined in terms of tableaux, *Advances in Applied Math.* **10** (1989), 396–438.

7.11. L. C. Biedenharn and J. D. Louck, Inhomogeneous basis set of symmetric polynomials defined by tableaux, *Proc. Nat. Acad. Sci. U.S.A.* **87** (1990), 1441–1445.

7.12. N. G. de Bruijn, Pólya's theory of counting, in *Applied Combinatorial Mathematics* (E. F. Beckenbach, ed.), Wiley, New York, 1964; reprinted by Krieger, Malabar, Florida, 1981.

7.13. W. H. Burge, Four correspondences between graphs and generalized Young tableaux, *J. Combinatorial Theory (A)* **17** (1974), 12–30.

7.14. W. Burnside, *Theory of Groups of Finite Order*, Cambridge University Press, 1897.

7.15. W. Burnside, *Theory of Groups of Finite Order*, second edition, Cambridge University Press, 1911; reprinted by Dover, New York, 1955.

7.16. L. Carlitz, Rectangular arrays and plane partitions, *Acta Arithmetica* **13** (1967), 29–47.

7.17. J. B. Carrell, Chern classes of the Grassmannians and the Schubert calculus, *Topology* **17** (1978), 177–182.

7.18. A. L. Cauchy, Mémoire sur les fonctions qui ne peuvent obtenir que deux valeurs égales et de signes contraires par suite des transpositions opérés entre les variables qu'elles renferment, *J. École Polyt.* **10** (1815), 29–112; *Oeuvres*, ser. 2, vol. 1, pp. 91–169.

7.19. A. L. Cauchy, Mémoire sur les fonctions alternées et sur les sommes alternées, *Exercices Anal. et Phys. Math.* **2** (1841), 151–159; *Oeuvres*, ser. 2, vol. 12, pp. 173–182.

7.20. A. L. Cauchy, Mémoire sur diverses propriétés remarquables des substitutions régulaires ou irrègulaires, et des systèmes de substitutions conjuguées (suite), *C. R. Acad. Sci. Paris* **21** (1845), 972–987; *Oeuvres*, ser. 1, vol. 9, pp. 371–387.

7.21. T. W. Chaundy, Partition-generating functions, *Quart. J. Math. (Oxford)* **2** (1931), 234–240.

7.22. W. Y. C. Chen and J. D. Louck, The factorial Schur function, *J. Math. Phys.* **34** (1993), 4144–4160.

7.23. H. Cohn, M. Larsen, and J. Propp, The shape of a typical boxed plane partition, preprint, 13 January 1998; available electronically at http://front.math.ucdavis.edu/math.CO/9801059.

7.24. C. W. Curtis, *Frobenius, Burnside, Schur and Brauer: Pioneers of Representation Theory*, American Mathematical Society, Providence, to appear.

7.25. S. Doty and G. Walker, Modular symmetric functions and irreducible modular representations of general linear groups, *J. Pure Appl. Algebra* **82** (1992), 1–26.

7.26. G. Duchamp, D. Krob, B. Leclerc, and J.-Y. Thibon, Fonctions quasi-symétriques, fonctions symétriques non commutatives et algèbres de Hecke à $q = 0$, *C. R. Acad. Sci. Paris Sér. I Math.* **322** (1996), 107–112.

7.27. G. Duchamp, A. Klyachko, D. Krob, and J.-Y. Thibon, Noncommutative symmetric functions III: Deformations of Cauchy and convolution algebras, LITP preprint 96/08, Paris, 1996.

7.28. C. Ehresmann, Sur la topologie de certains espaces homogènes, *Ann. Math.* **35** (1934), 396–443.
7.29. P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Math.* **2** (1935), 463–470.
7.30. W. Feit, The degree formula for the skew-representations of the symmetric group, *Proc. Amer. Math. Soc.* **4** (1953), 740–744.
7.31. S. V. Fomin, Two-dimensional growth in Dedekind lattices, M.S. thesis, Leningrad State University, 1979.
7.32. S. V. Fomin, Generalized Robinson–Schensted–Knuth correspondence (in Russian), *Zapiski Nauchn. Sem. LOMI* **155** (1986), 156–175.
7.33. S. Fomin, Duality of graded graphs, *J. Algebraic Combinatorics* **3** (1994), 357–404.
7.34. S. Fomin, Schensted algorithms for dual graded graphs, *J. Algebraic Combinatorics* **4** (1995), 5–45.
7.35. S. Fomin and C. Greene, Noncommutative Schur functions and their applications, *Discrete Math.*, to appear.
7.36. H. O. Foulkes, A survey of some combinatorial aspects of symmetric functions, in *Permutations (Actes Colloq., Univ. René-Descartes, Paris, 1972)*, Gauthiers-Villars, Paris, 1974, pp. 79–92.
7.37. H. O. Foulkes, Enumeration of permutations with prescribed up–down and inversion sequences, *Discrete Math.* **15** (1976), 235–252.
7.38. J. S. Frame, G. de B. Robinson, and R. M. Thrall, The hook graphs of S_n, *Canad. J. Math.* **6** (1954), 316–324.
7.39. D. S. Franzblau and D. Zeilberger, A bijective proof of the hook-length formula, *J. Algorithms* **3** (1982), 317–343.
7.40. F. G. Frobenius, Über die Congruenz nach einem aus zwei endlichen Gruppen gebildeten Doppelmodul, *J. Reine Angew. Math. (=Crelle's J.)* **101** (1887), 273–299; reprinted in *Gesammelte Abhandlungen*, vol. 2, Springer-Verlag, Heidelberg, 1988, pp. 304–330.
7.41. W. Fulton, *Intersection Theory*, Springer-Verlag, New York/Berlin, 1984.
7.42. W. Fulton and J. Harris, *Representation Theory*, Graduate Texts in Mathematics **129**, Springer-Verlag, New York/Berlin, 1991.
7.43. A. M. Garsia and M. Haiman, Some natural bigraded S_n-modules and q, t-Kostka coefficients, *Electron. J. Combinatorics* **3** (1996), R24, 60 pp.; also in *The Foata Festschrift* (J. Désarménien, A. Kerber, and V. Strehl, eds.), Imprimerie Louis-Jean, Gap, France, 1996, pp. 561–620.
7.44. I. M. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V. S. Retakh, and J.-Y. Thibon, Noncommutative symmetric functions, *Advances in Math.* **112** (1995), 218–348.
7.45. I. M. Gelfand and M. L. Tsetlin (= Cetlin), Finite-dimensional representations of the group of unimodular matrices (in Russian), *Doklady Akad. Nauk SSSR (N.S.)* **71** (1950), 825–828.
7.46. I. M. Gessel, Multipartite P-partitions and inner products of skew Schur functions, *Contemp. Math.* **34** (1984), 289–301.
7.47. I. M. Gessel and G. X. Viennot, Determinants, paths, and plane partitions, preprint, 28 July 1989.
7.48. G. Giambelli, Risoluzione del problema degli spazi secanti, *Mem. R. Accad. Sci. Torino (2)* **52** (1903), 171–211.
7.49. A. Girard, *Invention Nouvelle en l'Algèbre*, Amsterdam, 1629.
7.50. B. Gordon, Notes on plane partitions. V, *J. Combinatorial Theory (B)* **11** (1971), 157–168.
7.51. C. Greene, A. Nijenhuis, and H. S. Wilf, A probabilistic proof of a formula for the number of Young tableaux of a given shape, *Advances in Math.* **31** (1979), 104–109.

7.52. M. Haiman, On mixed insertion, symmetry, and shifted Young tableaux, *J. Combinatorial Theory (A)* **50** (1989), 196–225.
7.53. J. I. Hall, E. M. Palmer, and R. W. Robinson, Redfield's lost paper in a modern context, *J. Graph Theory* **8** (1984), 225–240.
7.54. P. Hall, The algebra of partitions, in *Proc. 4th Canadian Math. Congress (Banff)*, 1959, pp.147–159; reprinted in *Percy Alexander MacMahon: Collected Papers*, vol. 1 (G. E. Andrews, ed.), MIT Press, Cambridge, Massachusetts, 1978, and in P. Hall, *The Collected Works of Philip Hall*, Oxford University Press, Oxford/New York, 1988, pp. 465–477.
7.55. F. Harary and R. W. Robinson, The rediscovery of Redfield's papers, *J. Graph Theory* **8** (1984), 191–192.
7.56. T. Hawkins, The origins of the theory of group characters, *Arch. Hist. Exact Sci.* **7** (1970/71), 142–170.
7.57. T. Hawkins, Hypercomplex numbers, Lie groups, and the creation of group representation theory, *Arch. Hist. Exact Sci.* **8** (1971), 243–287.
7.58. T. Hawkins, New light on Frobenius' creation of the theory of group characters, *Arch. Hist. Exact Sci.* **12** (1974), 217–243.
7.59. T. Hawkins, The creation of the theory of group characters, *Rice Univ. Studies* **64** (1978), 57–71.
7.60. A. P. Hillman and R. M. Grassl, Reverse plane partitions and tableaux hook numbers, *J. Combinatorial Theory (A)* **21** (1976), 216–221.
7.61. W. V. D. Hodge, The basis for algebraic varieties of given dimension on grassmannian varieties, *J. London Math. Soc.* **16** (1941), 245–255.
7.62. W. V. D. Hodge, The intersection formulae for a grassmannian variety, *J. London Math. Soc.* **17** (1942), 48–64.
7.63. G. Horrocks, On the relation of S-functions to Schubert varieties, *Proc. London Math. Soc.* **7** (1957), 265–280.
7.64. K. G. J. Jacobi, De functionibus alternantibus earumque divisione per productum e differentiis elementorum conflatum, *J. Reine Angew Math.* (= *Crelle's J.*) **22** (1841), 360–371; *Werke* **3**, 439–452. A German translation was given by G. Stäckel, in *Über die Bildung und die Eigenschaften der Determinanten*, Leipzig, 1896, 73 pp.
7.65. G. D. James and A. Kerber, *The Representation Theory of the Symmetric Group*, Encyclopedia of Mathematics and its Applications **16**, Addison-Wesley, Reading, Massachusetts, 1981.
7.66. W. W. Johnson, On the calculation of the co-factor of alternants of the fourth order, *Amer. J. Math.* **7** (1885), 380–388.
7.67. A. Kerber, *Algebraic Combinatorics via Finite Group Actions*, Wissenschaftsverlag, Mannheim/Vienna/Zürich, 1991.
7.68. S. L. Kleiman, Problem 15. Rigourous foundation of Schubert's enumerative calculus, in *Mathematical Developments Arising From Hilbert Problems* (Proc. Symp. Pure Math. **19**), American Mathematical Society, Providence, RI, 1976, pp. 445–482.
7.69. S. L. Kleiman and D. Laksov, Schubert calculus, *Amer. Math. Monthly* **79** (1972), 1061–1082.
7.70. D. E. Knuth, A note on solid partitions, *Math. Comp.* **24** (1970), 955–961.
7.71. D. E. Knuth, Permutations, matrices, and generalized Young tableaux, *Pacific J. Math.* **34** (1970), 709–727.
7.72. D. E. Knuth, *The Art of Computer Programming*, vol. 3, *Sorting and Searching*, Addison-Wesley, Reading, Massachusetts, 1973.
7.73. C. Kostka, Über die Bestimmung von symmetrischen Funktionen der Wurzeln einer algebraischen Gleichung durch deren Coefficienten, *Crelle's J.* **81** (1875), 281–289.

7.74. C. Kostka, Über den Zusammenhang zwischen einigen Formen von symmetrischen Functionen, *Crelle's J.* **93** (1882), 89–123.
7.75. C. Kostka, Tafeln für symmetrische Funktionen bis zur elften Dimension, *Wissenschaftliche Beilage zum Programm des königl. Gymnasiums and Realgymnasiums zu Insterberg*, 1908.
7.76. C. Krattenthaler, Bijective proofs of the hook formulas for the number of standard Young tableaux, ordinary and shifted, *Electron. J. Combinatorics* **2** (1995), R13, 9pp.
7.77. C. Krattenthaler, An involution principle-free bijective proof of Stanley's hook-content formula, *Discrete Math. Theoret. Computer Science* (1998), to appear.
7.78. C. Krattenthaler, A bijective proof of the hook-content formula for super Schur functions and a modified jeu de taquin, *Electron. J. Combinatorics* **3** (1996), R14, 24 pp.; also in *The Foata Festschrift* (J. Désarménien, A. Kerber, and V. Strehl, eds.), Imprimerie Louis-Jean, Gap, France, 1996, pp. 373–396.
7.79. C. Krattenthaler, Another involution principle-free proof of Stanley's hook-content formula, preprint.
7.80. D. Krob, B. Leclerc, and J.-Y. Thibon, Noncommutative symmetric functions. II. Transformations of alphabets, *Internat. J. Algebra Comput.* **7** (1997), 181–264.
7.81. D. Krob and J.-Y. Thibon, Noncommutative symmetric functions IV: Quantum linear groups and Hecke algebras at $q = 0$, *J. Algebraic Combinatorics* **6** (1997), 339–376.
7.82. M. A. A. van Leeuwen, The Robinson–Schensted and Schützenberger algorithms, Part 1: new combinatorial proofs, Preprint AM-R9208 1992, Centrum voor Wiskunde en Informatica, 1992.
7.83. M. A. A. van Leeuwen, The Robinson–Schensted and Schützenberger algorithms, an elementary approach, *Electron. J. Combinatorics* **3**, R15 (1996), 46 pp.; reprinted in *The Foata Festschrift* (J. Désarménien, A. Kerber, and V. Strehl, eds.), Imprimerie Louis-Jean, Gap, 1996, pp. 397–442.
7.84. L. Lesieur, Les problèmes d'intersection sur une variété de Grassmann, *C. R. Acad. Sci. Paris* **225** (1947), 916–917.
7.85. D. E. Littlewood, The construction of invariant matrices, *Proc. London Math. Soc. (2)* **43** (1937), 226–240.
7.86. D. E. Littlewood, Invariant theory, tensors, and group characters, *Phil. Trans. Roy. Soc. London (A)* **239** (1944), 305–365.
7.87. D. E. Littlewood, *A University Algebra*, William Heineman Ltd., Melbourne/London/Toronto, 1950; second edition, 1958.
7.88. D. E. Littlewood, *The Theory of Group Characters and Matrix Representations of Groups*, second edition, Oxford University Press, London, 1950; reprinted 1958.
7.89. D. E. Littlewood and A. R. Richardson, Group characters and algebra, *Phil. Trans. Royal Soc. A (London)* **233** (1934), 99–141.
7.90. D. E. Littlewood and A. R. Richardson, Some special S-functions and q-series, *Quart. J. Math. (Oxford)* **6** (1935), 184–198.
7.91. E. K. Lloyd, J. Howard Redfield: 1879–1944, *J. Graph Theory* **8** (1984), 195–203.
7.92. I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, Oxford University Press, Oxford, 1979.
7.93. I. G. Macdonald, Polynomial functors and wreath products, *J. Pure Appl. Algebra* **18** (1980), 173–204.
7.94. I. G. Macdonald, The algebra of partitions, in *Group Theory*, Academic Press, London/New York, 1984, pp. 315–333.
7.95. I. G. Macdonald, *Notes on Schubert Polynomials*, Publ. LACIM **6**, Univ. du Québec à Montréal, 1991.
7.96. I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, second edition, Oxford University Press, Oxford, 1995.

7.97. I. G. Macdonald, Schur functions: themes and variations, in *Séminaire Lotharingien de Combinatoire*, Publ. I.R.M.A. Strasbourg **498**, 1992, pp. 5–39.

7.98. P. A. MacMahon, Memoir on the theory of the partitions of numbers – Part I, *Phil. Trans. Royal Soc. London (A)* **187** (1897), 619–673; *Collected Works*, vol. 1 (G. E. Andrews, ed.), MIT Press, Cambridge, Massachusetts, 1978, pp. 1026–1080.

7.99. P. A. MacMahon, Memoir on the theory of the partitions of numbers – Part IV, *Phil. Trans. Royal Soc. London (A)* **209** (1909), 153–175; *Collected Works*, vol. 1 (G. E. Andrews, ed.), MIT Press, Cambridge, Massachusetts, 1978, pp. 1292–1314.

7.100. P. A. MacMahon, Partitions of numbers whose graphs possess symmetry, *Trans. Cambridge Phil. Soc.* **17** (1899), 149–170; *Collected Works*, vol. 1 (G. E. Andrews, ed.), MIT Press, Cambridge, Massachusetts, 1978, pp. 1384–1405.

7.101. P. A. MacMahon, *Combinatory Analysis*, vols. 1 and 2, Cambridge University Press, Cambridge, 1915, 1916; reprinted in one volume by Chelsea, New York, 1960.

7.102. C. Malvenuto and C. Reutenauer, Duality between quasi-symmetric functions and the Solomon descent algebra, *J. Algebra* **177** (1995), 967–982.

7.103. O. H. Mitchell, Note on determinants of powers, *Amer. J. Math.* **4** (1882), 341–344.

7.104. A. Molev, Factorial supersymmetric Schur functions and super Capelli identities, in *Kirillov's Seminar on Representation Theory* (G. I. Olshanski, ed.), Amer. Math. Soc. Translations (2) **181**, American Mathematical Society, Providence, 1998, pp. 109–137.

7.105. T. Muir, On the quotient of a simple alternant by the difference-product of the variable, *Proc. Royal Soc. Edinburgh* **14** (1887), 433–445.

7.106. T. Muir, *The Theory of Determinants in the Historical Order of Development*, vol. 1, second edtion, MacMillan, London, 1906.

7.107. F. D. Murnaghan, On representations of the symmetric group, *Amer. J. Math.* **59** (1937), 437–488.

7.108. H. Nägelsbach, Ueber eine Classe symmetrischer Functionen, *Sch.-Prog.*, Zweibrücken, 1871, 42 pp.

7.109. T. Nakayama, On some modular properties of irreducible representations of a symmetric group, I., *Japan. J. Math.* **17** (1941), 165–184.

7.110. V. S. Nanda, Partition theory and thermodynamics of multi-dimensional harmonic oscillator assemblies, *Proc. Camb. Phil. Soc.* **47** (1951), 591–601.

7.111. V. S. Nanda, Tables of solid partitions, *Proc. Nat. Inst. Sci. India (A)* **19** (1953), 313–314.

7.112. P. M. Neumann, A lemma that is not Burnside's, *Math. Sci.* **4** (1979), 133–141.

7.113. J.-C. Novelli, I. Pak, and A. V. Stoyanovskii, A new proof of the hook-length formula, *Discrete Math. and Theoretical Computer Sci.* **1** (1997), 053–067; available electronically at http://www.compscinet.com/dm/dm0101ind.html.

7.114. S. Okada, Wreath products by the symmetric groups and product posets of Young's lattices, *J. Combinatorial Theory (A)* **55** (1990), 14–32.

7.115. A. Okounkov, Binomial formula for Macdonald polynomials, *Math. Res. Letters*. to appear; available electronically at http://xxx.lanl.gov/list/q-alg/9608.

7.116. A. Okounkov and G. Olshanski, Shifted Schur functions, *Algebra i Analiz* **9** (1997), 73–146 (Russian); English transtation to appear in *St.Petersburg Math. J.* **9** (1998); available electronically at http://xxx.lanl.gov/list/q-alg/9605.

7.117. A. Okounkov and G. Olshanski, Shifted Schur functions II. The binomial formula for characters of classical groups and its applications, in *Kirillov's Seminar on Representation Theory* (G. I. Olshanski, ed.), Amer. Math. Soc. Translations (2) **181**, American Mathematical Society, Providence, 1998, pp. 245–271.

7.118. M. Pieri, Sul problema degli spazi secanti. Nota 1^{a}, *Rend. Ist. Lombardo (2)* **26** (1893), 534–546.

7.119. G. Pólya, Kombinatorische Anzahlbestimmungen für Gruppen, Graphen, und chemische Verbindungen, *Acta Math.* **68** (1937), 145–254.
7.120. G. Pólya and R. C. Read, *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds*, Springer-Verlag, New York/Berlin, 1987.
7.121. R. A. Proctor, Bruhat lattices, plane partition generating functions, and minuscule representations, *Europ. J. Combinatorics* **5** (1984), 331–350.
7.122. S. Rabinowitz, Advanced Problem 5641, *Amer. Math. Monthly* **75** (1968), 1125.
7.123. R. C. Read, The use of S-functions in combinatorial analysis, *Canad. J. Math.* **20** (1968), 808–841.
7.124. J. H. Redfield, The theory of group reduced distributions, *Amer. J. Math.* **49** (1927), 433–455.
7.125. J. H. Redfield, Enumeration by frame group and range groups, *J. Graph Theory* **8** (1984), 205–223.
7.126. V. Reiner and M. Shimozono, Key polynomials and a flagged Littlewood–Richardson rule, *J. Combinatorial Theory (A)* **70** (1995), 107–143.
7.127. J. B. Remmel, Bijective proofs of formulae for the number of standard Young tableaux, *Linear and Multilinear Algebra* **11** (1982), 45–100.
7.128. J. B. Remmel and R. Whitney, A bijective proof of the hook formula for the number of column strict tableaux with bounded entries, *Europ. J. Combinatorics* **4** (1983), 45–63.
7.129. J. B. Remmel and R. Whitney, A bijective proof of the generating function for the number of reverse plane partitions via lattice paths, *Linear and Multilinear Algebra* **16** (1984), 75–91.
7.130. C. Reutenauer, *Free Lie Algebras*, Oxford University Press, Oxford, 1993.
7.131. G. de B. Robinson, On the representations of S_n, *Amer. J. Math.* **60** (1938), 745–760.
7.132. T. Roby, *Applications and Extensions of Fomin's Generalization of the Robinson–Schensted Correspondence to Differential posets*, Ph.D. thesis, Massachusetts Institute of Technology, 1991.
7.133. T. Roby, The connection between the Robinson–Schensted correspondence for skew oscillating tableaux and graded graphs, *Discrete Math.* **139** (1995), 481–485.
7.134. B. E. Sagan, The ubiquitous Young tableau, in *Invariant theory and tableaux (Minneapolis, MN, 1988)* (D. Stanton, ed.), IMA Vol. Math. Appl. **19**, Springer-Verlag, New York/Berlin, 1990, pp. 262–298.
7.135. B. E. Sagan, *The Symmetric Group*, Wadsworth & Brooks/Cole, Pacific Grove, CA, 1991.
7.136. C. E. Schensted, Longest increasing and decreasing subsequences, *Canad. J. Math.* **13** (1961), 179–191.
7.137. I. Schur, Über eine Klasse von Matrizen die sich einer gegebenen Matrix zuordnen lassen, Inaugural–Dissertation, Friedrich-Wilhelms-Universität zu Berlin, 1901; *Ges. Abhandlungen*, vol. 1, pp. 1–72.
7.138. I. Schur, Aufgabe 569, *Arch. Math. Phys. (3)* **27**, (1918), 163; *Ges. Abhandlungen*, vol. 3, p. 456.
7.139. I. Schur, Über die rationalen Darstellungen der allgemeinen linearen Gruppe, *Sitzungsber. Preuss. Akad. Wissenschaften* **10** (1927) 58–75; *Ges. Abhandlungen*, vol. 3, pp. 68–85.
7.140. M. P. Schützenberger, Quelques remarques sur une construction de Schensted, *Math. Scand.* **12** (1963), 117–128.
7.141. A. Seidenberg, A simple proof of a theorem of Erdős and Szekeres, *J. London Math. Soc.* **34** (1959), 352.
7.142. J.-P. Serre, *Linear Representations of Finite Groups*, Springer-Verlag, New York/Heidelberg/Berlin, 1977.

7.143. W. Specht, Eine Verallgemeinerung der symmetrischen Gruppe, *Schriften Math. Seminar Berlin* **1** (1932), 1–32.

7.144. R. Stanley, Solution to Advanced Problem 5641, *Amer. Math. Monthly* **76** (1969), 1153.

7.145. R. Stanley, Theory and application of plane partitions, Part 1, *Studies in Applied Math.* **50** (1971), 167–188.

7.146. R. Stanley, Theory and application of plane partitions, Part 2, *Studies in Applied Math.* **50** (1971), 259–279.

7.147. R. Stanley, The conjugate trace and trace of a plane partition, *J. Combinatorial Theory* **14** (1973), 53–65.

7.148. R. Stanley, Some combinatorial aspects of the Schubert calculus, in *Combinatoire et Répresentation du Groupe Symétrique (Strasbourg, 1976)*, Lecture Notes in Math. **579**, Springer-Verlag, Berlin, 1977, pp. 217–251.

7.149. R. Stanley, Invariants of finite groups and their applications to combinatorics, *Bull (N. S.) Amer. Math. Soc.* **1** (1979), 475–511.

7.150. R. Stanley, Some combinatorial properties of Jack symmetric functions, *Advances in Math.* **77** (1989), 76–115.

7.151. D. Stanton and D. White, *Constructive Combinatorics*, Springer-Verlag, New York/Berlin, 1986.

7.152. S. Sundaram, Tableaux in the representation theory of the classical Lie groups, in *Invariant Theory and Tableaux (Minneapolis, MN, 1988)*, IMA Vol. Math. Appl. **19**, Springer-Verlag, New York, 1990, pp. 191–225.

7.153. N. Trudi, Intorno ad un determinante piu generale di quello che suol dirsi determinante delle radici di una equazione, ed alle funzioni simmetriche complete di questi radici, *Rendic. dell' Accac. (Napoli)* (1864), 121–134; with a slightly different title in *Giornale di Mat.* **2** (1864), 152–158, 180–186.

7.154. K. Th. Vahlen, IB3b. Rationale Funktionen der Wurzeln; symmetrische und Affektfunktionen, in *Encyklopädie der Mathematischen Wissenschaften*, Erster Band, Teubner, Leipzig, 1898–1904, pp. 449–479.

7.155. G. Viennot, Une forme géometrique de la correspondance de Robinson–Schensted in *Combinatoire et représentation du groupe symétrique (Strasbourg, 1976)*, Lecture Notes in Math. **579**, Springer-Verlag, Berlin/Heidelberg/New York, 1977. pp. 29–58.

7.156. G. Viennot, Chain and antichain families, grids and Young tableaux, in *Orders: Descriptions and Roles (L'Arbresle, 1982)*, North-Holland Math. Stud. **99**, North-Holland, Amsterdam/New York, 1984, pp. 409–463.

7.157. M. L. Wachs, Flagged Schur functions, Schubert polynomials, and symmetrizing operators, *J. Combinatorial Theory (A)* **40** (1985), 276–289.

7.158. B. L. van der Waerden, Topologische Begründung des Kalküls der abzählenden Geometrie, *Math. Ann.* **102** (1930), 337–362.

7.159. H. Weyl, *The Classical Groups*, second edition, Princeton University Press, Princeton, New Jersey, 1946.

7.160. E. M. Wright, Rotatable partitions, *J. London Math. Soc.* **43** (1968), 501–505.

7.161. E. M. Wright, Burnside's lemma: A historical note, *J. Combinatorial Theory (B)* **30** (1981), 89–90.

7.162. A. Young, Qualitative substitutional analysis (third paper), *Proc. London Math. Soc. (2)* **28** (1927), 255–292.

7.163. A. V. Zelevinsky, *Representations of Finite Classical Groups, a Hopf Algebra Approach*, Lecture Notes in Math. **569**, Springer-Verlag, Berlin/Heidelberg/New York, 1981.

Chapter 7: Appendix 1

Knuth Equivalence, Jeu de Taquin, and the Littlewood–Richardson Rule

(by Sergey Fomin)

This Appendix is devoted to the study of several combinatorial constructions involving standard Young tableaux (SYTs) that lead to the proof of the Littlewood–Richardson rule, a combinatorial rule describing the coefficients in the Schur function expansion of an arbitrary skew Schur function (or in a product of two ordinary Schur functions).

Most of what follows can be straightforwardly generalized to semistandard Young tableaux (SSYTs). We do not do it here, in order to simplify the presentation.

A1.1 Knuth Equivalence and Greene's Theorem

The RSK algorithm $w \xrightarrow{\text{RSK}} (P, Q)$ associates to a permutation $w \in \mathfrak{S}_n$ a pair of SYTs: the *insertion tableau* P and the *recording tableau* Q; these tableaux have the same shape $\mathrm{sh}(w)$. In this section, we examine the following two questions:

- What are the conditions for two permutations to have the same shape $\mathrm{sh}(w)$?
- What are the conditions for two permutations to have the same insertion tableau P?

The first question has an answer involving a particular family of poset-theoretic invariants of permutations. The equivalence relation appearing in the second question can be described in terms of certain elementary transformations that change three consecutive entries of a permutation. We first state these two results, and devote the rest of this section to their proof.

For a permutation $w = w_1 \cdots w_n \in \mathfrak{S}_n$ and $k \in \mathbb{N}$, let $I_k = I_k(w)$ denote the maximal number of elements in a union of k increasing subsequences of w. Analogously, let D_l be the maximal size of a union of l decreasing subsequences of w. For example, for $w = 236145 \in \mathfrak{S}_6$, we have: $I_0 = 0$, $I_1 = 4$, $I_2 = I_3 = \cdots = 6$; $D_0 = 0$, $D_1 = 2$, $D_2 = 4$, $D_3 = 5$, $D_4 = D_5 = \cdots = 6$.

A1.1.1 Theorem (Greene's theorem). *Let* $w \in \mathfrak{S}_n$ *and* $\mathrm{sh}(w) = \lambda$. *Then, for any positive integer* k *and* l,

$$\begin{aligned} I_k(w) &= \lambda_1 + \cdots + \lambda_k\,, \\ D_l(w) &= \lambda_1' + \cdots + \lambda_l'\,. \end{aligned} \tag{A1.123}$$

(Note that Theorem A1.1.1 is a restatement of Theorems 7.23.13 and 7.23.17.)

To illustrate, take $w = 236145$. Then $w \xrightarrow{\mathrm{RSK}} (P, Q)$ with

$$P = \begin{array}{|c|c|c|c|}\hline 1 & 3 & 4 & 5 \\ \hline 2 & 6 \\ \cline{1-2} \end{array}\,, \qquad Q = \begin{array}{|c|c|c|c|}\hline 1 & 2 & 3 & 6 \\ \hline 4 & 5 \\ \cline{1-2} \end{array}\,, \qquad \mathrm{sh}(w) = \lambda = (4, 2)\,.$$

To obtain the numbers I_k, we count boxes in the first several rows of the shape: $0, 4, 6, 6, \ldots$. Analogously, counting boxes in the first several columns of λ gives $0, 2, 4, 5, 6, 6, \ldots$, agreeing with our previous computations.

Theorem A1.1.1 implies that two permutations have the same shape $\mathrm{sh}(w)$ if and only if the values $I_1, I_2, \ldots$ (or $D_1, D_2, \ldots$) computed for these permutations are the same. Another direct implication of Theorem A1.1.1 is given below.

A1.1.2 Corollary. *For any permutation* w, *the sequences* $(I_1, I_2 - I_1, I_3 - I_2, \ldots)$ *and* $(D_1, D_2 - D_1, D_3 - D_2, \ldots)$ *define conjugate partitions.*

To formulate an answer to the second question posed at the beginning of this section, we will need the following definition.

A1.1.3 Definition. A *Knuth transformation* of a permutation is its transformation into another permutation that has one of the following forms:

$$\begin{array}{cccc} \cdots acb \cdots & \cdots cab \cdots & \cdots bac \cdots & \cdots bca \cdots \\ \downarrow & \downarrow & \downarrow & \downarrow \\ \cdots cab \cdots & \cdots acb \cdots & \cdots bca \cdots & \cdots bac \cdots \end{array} \tag{A1.124}$$

where $a < b < c$ (all other entries remain intact). Thus each Knuth transformation switches two adjacent entries a and c provided an entry b satisfying $a < b < c$ is located next to a or c. Two permutations $u, v \in \mathfrak{S}_n$ are called *Knuth-equivalent* (denoted $u \overset{K}{\sim} v$) if one of them can be obtained from another by a sequence of Knuth transformations.

For example, the six permutations in (A1.125) below form a Knuth equivalence class; the ones that differ by a single Knuth transformation are connected by

an edge.

$$
\begin{array}{ccccc}
 & 51243 & \text{———} & 15243 & \text{———} & 12543 \\
 & | & & | & & \\
54123 \;\text{———} & 51423 & \text{———} & 15423 & &
\end{array}
\tag{A1.125}
$$

A1.1.4 Theorem. *Permutations are Knuth-equivalent if and only if their insertion tableaux coincide.*

Permutations u and v are said to be *dual Knuth-equivalent* if $u^{-1} \overset{K}{\sim} v^{-1}$. For instance, 34521 is dual Knuth-equivalent to 12543. Recall the following symmetry of the RSK algorithm (see Theorem 7.13.1): the recording tableau for a permutation w is nothing but the insertion tableau for w^{-1}. Thus Theorem A1.1.4 implies that permutations have the same *recording* tableaux if and only if they are dual Knuth-equivalent.

Knuth equivalence classes can be given a more detailed description, which is provided in Theorem A1.1.6 below.

A1.1.5 Definition. Let T be a tableau. The *reading word* of T (denoted reading (T)) is the sequence of entries of T obtained by concatenating the rows of T bottom to top. For example, the tableau

$$
\begin{array}{|c|c|c|c|}
\cline{3-4}
\multicolumn{1}{c}{} & \multicolumn{1}{c|}{} & 1 & 2 \\
\hline
3 & 5 & 6 & 8 \\
\hline
4 & 7 & 9 & \multicolumn{1}{c}{} \\
\cline{1-3}
\end{array}
$$

has the reading word 479356812.

In what follows, we say that a tableau has a *straight shape* if its shape is a Young (or Ferrers) diagram. Observe that any straight-shape tableau T is uniquely reconstructed from its reading word. Indeed, to break a word $w = \text{reading}(T)$ into segments representing the rows of T, simply locate the descents of w.

A1.1.6 Theorem. *Each Knuth equivalence class contains exactly one reading word of a straight-shape SYT (call this tableau T), and consists of all permutations whose insertion tableau is T.*

For example, the only reading word in the Knuth equivalence class shown in (A1.125) is

$$
54123 = \text{reading}(T)\,, \qquad T = \begin{array}{|c|c|c|}
\hline
1 & 2 & 3 \\
\hline
4 & \multicolumn{2}{c}{} \\
\cline{1-1}
5 & \multicolumn{2}{c}{} \\
\cline{1-1}
\end{array}\,.
$$

There are indeed six permutations with insertion tableau T (since $f^{\mathrm{sh}(T)} = 6$), and these are exactly the ones appearing in (A1.125).

Proofs of Theorems A1.1.1, A1.1.4, and A1.1.6

A1.1.7 Lemma. *For any k, the values $I_k(w)$ and $D_k(w)$ are invariant under Knuth transformations of a permutation w.*

Proof. It is enough to prove the invariance of the numbers I_k, since replacing a permutation $w = w_1 \cdots w_n$ by $w' = w_n \cdots w_1$ interchanges I_k and D_k, while clearly $u \overset{K}{\sim} v \Leftrightarrow u' \overset{K}{\sim} v'$. We need to show that I_k does not change under each of the two types of Knuth transformations:

$$u = \cdots acb \cdots \to v = \cdots cab \cdots, \qquad a < b < c,$$

and

$$u = \cdots bac \cdots \to v = \cdots bca \cdots, \qquad a < b < c$$

(cf. (A1.124)). Since these two cases are completely analogous, let us concentrate on the first one. Let $I_k(u) = m$. Obviously, $I_k(v) \leq m$. Moreover, the only situation where we may possibly have $I_k(v) < m$ is the following: *every* collection $\{\sigma_1, \ldots, \sigma_k\}$ of k disjoint increasing subsequences of u which jointly cover m elements has an element (say, σ_1) containing both a and c. Suppose this situation does indeed take place, and consider such a collection $\{\sigma_1, \ldots, \sigma_k\}$. If b does not belong to any σ_i, then simply replace c by b in σ_1, arriving at a contradiction with our assumption. We thus may assume that b belongs, say, to σ_2:

$$\sigma_1 = \left(u_{i_1} < \cdots < u_{i_s} < a < c < u_{i_{s+3}} < \cdots\right)$$

$$\sigma_2 = \left(u_{j_1} < \cdots < u_{j_t} < b < u_{j_{t+2}} < \cdots\right).$$

Then the increasing subsequences σ_1' and σ_2' defined by

$$\sigma_1' = (u_{i_1} < \cdots < u_{i_s} < a < b < u_{i_{t+2}} < \cdots)$$

$$\sigma_2' = (u_{j_1} < \cdots < u_{j_t} < c < u_{i_{s+3}} < \cdots)$$

will jointly cover the same elements of u as σ_1 and σ_2 do. The collection $\{\sigma_1', \sigma_2', \sigma_3, \ldots, \sigma_k\}$ will cover m elements, while not containing a subsequence to which both a and c belong. This contradicts our assumption, and the proof follows. □

We next show that the RSK insertion algorithm can be viewed as a sequence of Knuth transformations.

A1.1.8 Lemma. *Any permutation is Knuth-equivalent to the reading word of its insertion tableau.*

Proof. Recall from Section 7.11 that $P \leftarrow k$ denotes the result of inserting k into P. To prove the lemma, it suffices to show that, for any (straight-shape) SYT P and any positive integer k, we have

$$\text{reading}(P) \cdot k \overset{K}{\sim} \text{reading}(P \leftarrow k)\,, \tag{A1.126}$$

where $\cdot$ stands for concatenation. Because of the row-by-row nature of the RSK insertion algorithm, it is enough to check (A1.126) for a single-row tableau. This is a straightforward verification. □

A1.1.9 Corollary. *Let P be the insertion tableau for w. Then the permutations w and* reading(P) *have the same values of parameters I_k and D_k, for all k.*

Proof. Directly follows from Lemmas A1.1.7 and A1.1.8. □

A1.1.10 Lemma. *Let w be the reading word of a straight-shape SYT T. Then T is the insertion tableau for w.*

Proof. In the special case of a tableau word, the RSK insertion process is very simple: increasing segments of the word are consecutively placed atop each other, eventually forming the original tableau. □

Proof of Theorem A1.1.1. In view of Corollary A1.1.9 and Lemma A1.1.10, we may assume that w is a reading word of a straight-shape SYT T. To illustrate, let

$$T = \begin{array}{|c|c|c|c|} \hline 1 & 3 & 4 & 8 \\ \hline 2 & 6 & 7 \\ \cline{1-3} 5 & 9 \\ \cline{1-2} \end{array} \quad ; \tag{A1.127}$$

then $w = 592671348$. Note that each row of T becomes an increasing subsequence in $w = \text{reading}(T)$. Thus, for any k,

$$I_k(w) \geq \lambda_1 + \cdots + \lambda_k\,. \tag{A1.128}$$

Furthermore, the entries of each column of T form a decreasing subsequence in w. Therefore, for any l,

$$D_l(w) \geq \lambda'_1 + \cdots + \lambda'_l\,. \tag{A1.129}$$

Let us now consider a box lying at the border of the shape λ (such as the boxes containing the entries $5, 9, 6, 7, 4, 8$ in the example (A1.127)). Assume that this box is located in row k and column l. Then

$$(\lambda_1 + \cdots + \lambda_k) + (\lambda'_1 + \cdots + \lambda'_l) = n + kl . \qquad \text{(A1.130)}$$

Combining (A1.128), (A1.129), and (A1.130), we obtain

$$I_k(w) + D_l(w) \geq n + kl .$$

On the other hand, an increasing and a decreasing subsequences may have at most one element in common. Hence

$$I_k(w) + D_l(w) \leq n + kl$$

for any k and l. Comparing this with the previous inequality, we conclude that $I_k(w) + D_l(w) = n + kl$, and moreover both (A1.128) and (A1.129) are actually *equalities* for the chosen values of k and l. Since every row and every column of λ contain at least one box that lies on the border, the identities (A1.123) hold for any k and l. □

A1.1.11 Corollary. *The shape* $\mathrm{sh}(w)$ *is invariant under Knuth transformations.*

Proof. In view of Theorem A1.1.1, the shape $\mathrm{sh}(w)$ is uniquely determined by the values $I_1(w), I_2(w), \ldots$, so the claim follows by Lemma A1.1.7. □

We will now show that a much stronger result holds.

A1.1.12 Corollary. *The insertion tableau* P *of a permutation* w *is invariant under Knuth transformations of* w.

Proof. For $k = 1, \ldots, n$, let $w_{(k)}$ denote the permutation in $\mathfrak{S}_k$ formed by the entries $1, \ldots, k$ of w. (For example, if $w = 236145$ then $w_{(4)} = 2314$.) Let us write $w \xrightarrow{\text{RSK}} (P, Q)$ and $w_{(k)} \xrightarrow{\text{RSK}} (P_{(k)}, Q_{(k)})$. Then $P_{(k)}$ is nothing but the tableau formed by the k smallest entries of P, since the larger entries do not interfere with the part of the insertion process that involves smaller entries. Now the crucial observation is the following: any Knuth transformation of w either does not change $w_{(k)}$ or else transforms the latter into a Knuth-equivalent permutation. By Corollary A1.1.11, this does not affect the shape $\mathrm{sh}(w_{(k)}) = \mathrm{sh}(P_{(k)})$. Since the tableau P can be viewed as a sequence of shapes

$$\emptyset \subset \mathrm{sh}(P_{(1)}) \subset \mathrm{sh}(P_{(2)}) \subset \cdots \subset \mathrm{sh}(P_{(n)}) = \mathrm{sh}(P) ,$$

and since all these shapes are unchanged by Knuth transformations, the proof follows. □

Proof of Theorems A1.1.4 and A1.1.6. It follows from Lemma A1.1.8 and Corollary A1.1.12 that two permutations are Knuth equivalent if and only if they have the same insertion tableau (whose reading word also belongs to the same equivalence class). Finally, two distinct reading words may not be Knuth-equivalent, since by Lemma A1.1.10 they have different insertion tableaux. □

A1.2 Jeu de Taquin

The constructions of the previous section are intimately related to the remarkable *jeu de taquin* equivalence relation among skew tableaux. In this section, we establish the fundamental properties of this equivalence, which will then be used in Section A1.3 to prove the main result – the Littlewood–Richardson rule.

Jeu de taquin, or the "teasing game," is a particular set of rules for transforming skew tableaux by viewing their entries as separate pieces that can be moved around on the "checkerboard" of the coordinate plane. These rules are designed so that the property of being a tableau was preserved. The concept of jeu de taquin is intuitively quite simple, and is probably best understood by looking at concrete examples. Still, we begin with a formal description.

A1.2.1 Definition. Let λ/μ be a skew shape. (In Figure A1-10, λ/μ is composed of the boxes made of solid lines segments.) Consider the boxes b that can be added to λ/μ, so that b shares at least one edge with λ/μ, and $\{b\} \cup \lambda/\mu$ is a valid skew shape. (In Figure A1-10, these boxes are made of dotted line segments.)

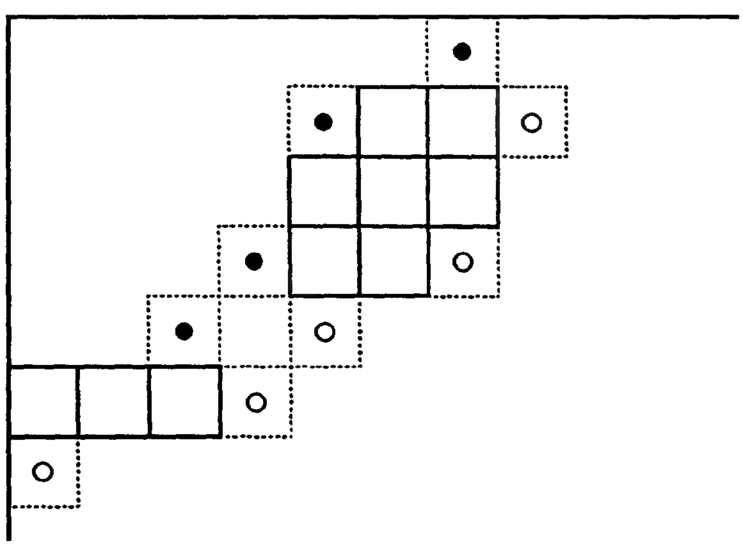


Figure A1-10. Adding boxes to a skew shape.

Two types of such boxes may occur, depending on the side of λ/μ that they are on. We mark by a bullet $\bullet$ the dashed boxes that share a lower or a right edge with λ/μ, while those that share an upper or a left edge are marked by a circle $\circ$.

Suppose we are given an SYT T of shape λ/μ. To each box b marked $\bullet$ or $\circ$, we will associate a transformation $\mathrm{jdt}_b(T)$ of T called a *jeu de taquin slide* of T into b. The definitions of the slides into inner boxes marked $\bullet$ and the outer boxes marked $\circ$ are completely analogous, so we will only discuss the first of these cases, and then provide examples illustrating both of them. Thus let us consider a box b_0 marked $\bullet$. There is at least one box b_1 in λ/μ that is adjacent to b_0 (i.e., such that b_0 and b_1 share an edge); if there are two such boxes, then let b_1 be the one with a smaller entry. Move the entry occupying b_1 into b_0. Then look at the tableau entries to the right and below b_1, and repeat the same procedure: if there is a unique such entry, then move it into b_1; if there are two to choose from, then move the smaller one. This will vacate some box b_2, and the process will continue until it reaches the outer boundary. The resulting tableau (indeed, it will be an SYT) is $\mathrm{jdt}_b(T)$, by definition.

For example, take

$$T = \begin{array}{ccccc} a & \boxed{1} & \boxed{3} & \boxed{7} & \boxed{10} \\ \boxed{2} & \boxed{5} & \boxed{6} & \boxed{9} & \\ \boxed{4} & \boxed{8} & \boxed{11} & b & \end{array} \tag{A1.131}$$

(the boxes a and b are not included in T). Then

$$\mathrm{jdt}_a(T) = \begin{array}{ccccc} \boxed{1} & \boxed{3} & \boxed{6} & \boxed{7} & \boxed{10} \\ \boxed{2} & \boxed{5} & \boxed{9} & & \\ \boxed{4} & \boxed{8} & \boxed{11} & & \end{array} \quad \text{and} \quad \mathrm{jdt}_b(T) = \begin{array}{ccccc} & \boxed{1} & \boxed{3} & \boxed{7} & \boxed{10} \\ & \boxed{2} & \boxed{6} & \boxed{9} & \\ \boxed{4} & \boxed{5} & \boxed{8} & \boxed{11} & \end{array} \tag{A1.132}$$

A1.2.2 Definition. Tableaux T and T' are called *jeu de taquin equivalent* (denoted $T \overset{\mathrm{jdt}}{\sim} T'$) if one can be obtained from another by a sequence of jeu de taquin slides.

Note that $\overset{\mathrm{jdt}}{\sim}$ is a symmetric (and obviously transitive) relation, since any jeu de taquin slide can be reversed by performing a slide into the box that was vacated at the previous stage. For instance, in the example (A1.131) we have $\mathrm{jdt}_c(\mathrm{jdt}_a(T)) = T$, where c is the box occupied by 9 in T.

A1.2.3 Lemma. *Each jeu de taquin slide converts the reading word of a tableau into a Knuth-equivalent one:* $\mathrm{reading}(\mathrm{jdt}_b(T)) \overset{K}{\sim} \mathrm{reading}(T)$.

Proof. Let us verify that at every step of the sliding process, the reading word is transformed into a Knuth-equivalent one. The horizontal slides do not change the reading word at all. A vertical slide of the form

$$\begin{array}{|c|c|c|c|c|c|c|c|c|}\hline a & \cdots & b & & c & \cdots & d & e & \cdots \\ \hline i & \cdots & j & k & l & \cdots & m \\ \cline{1-7}\end{array} \longrightarrow \begin{array}{|c|c|c|c|c|c|c|c|c|}\hline a & \cdots & b & k & c & \cdots & d & e & \cdots \\ \hline i & \cdots & j & & l & \cdots & m \\ \cline{1-7}\end{array}$$

replaces the segment

$$i \cdots jkl \cdots ma \cdots bc \cdots d$$

of the reading word by the segment

$$i \cdots jl \cdots ma \cdots bkc \cdots d.$$

To show that these two segments are Knuth-equivalent, we may use Theorem A1.1.4. Indeed, one easily checks that both segments have insertion tableau

$$\begin{array}{|c|c|c|c|c|c|c|}\hline a & \cdots & b & k & c & \cdots & d \\ \hline i & \cdots & j & l & \cdots & m \\ \cline{1-6}\end{array}\,,$$

and the lemma follows. □

The following result is sometimes called "the fundamental theorem of jeu de taquin."

A1.2.4 Theorem. *Each jeu de taquin equivalence class contains exactly one straight-shape tableau.*

Proof. If T is a tableau of a skew shape λ/μ, then performing consecutive slides into all boxes of μ (in any allowable order) will result in a straight-shape tableau, which is jeu de taquin equivalent to T. The uniqueness of such representative in a given equivalence class follows directly from Lemma A1.2.3 and the second statement of Theorem A1.1.4. □

We will use the notation $\mathrm{jdt}(T)$ to denote the unique straight-shape tableau in the jeu de taquin equivalence class of a given (skew) tableau T. Note that Lemma A1.2.3 implies that

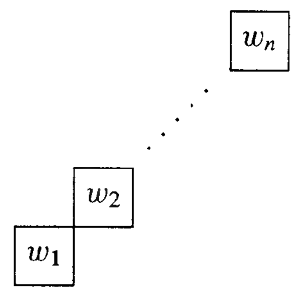
$$\mathrm{reading}(\mathrm{jdt}(T)) \overset{K}{\sim} \mathrm{reading}(T)\,. \tag{A1.133}$$

A1.2.5 Corollary. *Two standard tableaux are jeu de taquin equivalent if and only if their reading words are Knuth-equivalent.*

Proof. Let T and T' be standard tableaux. Then

$$\begin{aligned} T \overset{\text{jdt}}{\sim} T' &\iff \mathrm{jdt}(T) = \mathrm{jdt}(T') && \text{(by Theorem A1.2.4)} \\ &\iff \mathrm{reading}(\mathrm{jdt}(T)) \overset{K}{\sim} \mathrm{reading}(\mathrm{jdt}(T')) && \text{(by Theorem A1.1.4)} \\ &\iff \mathrm{reading}(T) \overset{K}{\sim} \mathrm{reading}(T') && \text{(by (A1.133)).} \end{aligned}$$ □

A1.2.6 Corollary. *For a permutation* $w = w_1 \cdots w_n \in \mathfrak{S}_n$, *let* T_w *denote the skew tableau*



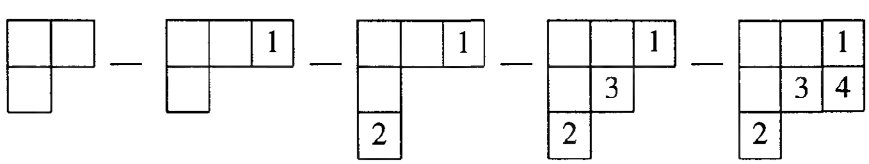
Then $\mathrm{jdt}(T_w)$ *is the insertion tableau for* w.

Proof. By Lemma A1.2.3, $\mathrm{reading}(\mathrm{jdt}(T_w)) \overset{K}{\sim} \mathrm{reading}(T_w) = w$, and the corollary follows from Theorem A1.1.4. □

Similarly to the RSK algorithm, jeu de taquin can be described in terms of *growth diagrams* (cf. Section 7.13). This is best explained by an example. The tableau

$$T = \begin{array}{ccc} & & \boxed{1} \\ & \boxed{3} & \boxed{4} \\ \boxed{2} & & \end{array} \qquad \text{(A1.134)}$$

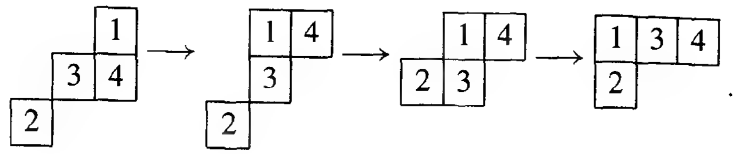
can be viewed as a sequence of shapes:



(disregard the entries, which are only shown to make the rules transparent), or as a sequence of partitions:

$$21 \;—31\; —311\; —321\; —331\,. \qquad \text{(A1.135)}$$

Consider the sequence of jeu de taquin slides



Replace each of these tableaux by the corresponding sequence of partitions, place these sequences on top of each other and rotate the resulting table to obtain the growth diagram shown in Figure A1-11.

Its upper left row (or perhaps it should be called column) corresponds to the original tableau T (cf. (A1.134)–(A1.135)). The lower right row is the tableau jdt(T) obtained as a result of this sequence of slides, and the lower left row

$$\emptyset — 1 — 11 — 21$$

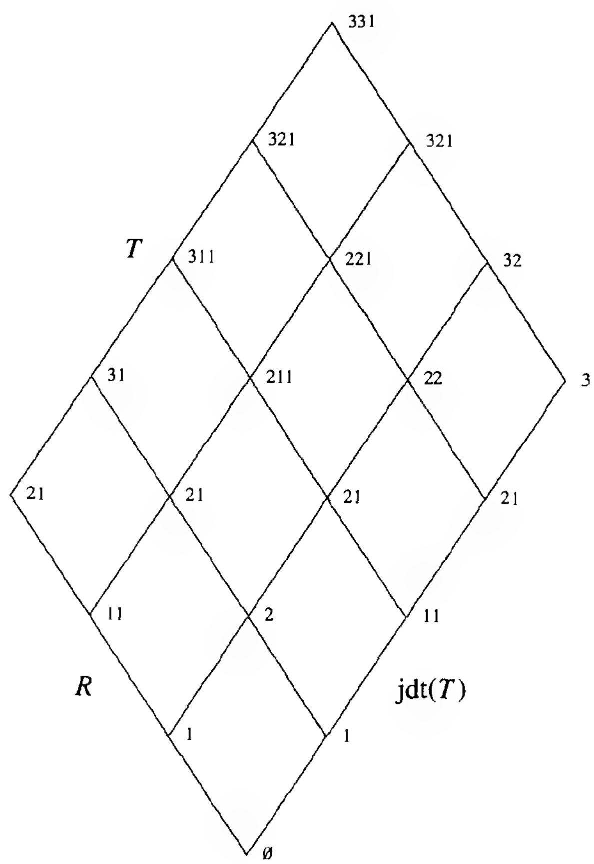


Figure A1-11. Growth diagram for jeu de taquin.

encodes the tableau

$$R = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & \multicolumn{1}{c}{} \\ \cline{1-1} \end{array},$$

which records the order in which the slides were performed: we first made a slide into the box occupied by 3, then into the box occupied by 2, and finally, into the one occupied by 1.

Note that by virtue of Theorem A1.2.4, the resulting tableau jdt(T) does not depend on the order in which the slides were performed, i.e., it does not depend on the tableau R.

Growth diagrams for sequences of jeu de taquin slides can be described by very simple local rules. First of all, it is easy to verify that whenever shape λ covers μ in a growth diagram, λ can be obtained by from μ by adding a single box. Another property of jeu de taquin growth is stated below.

A1.2.7 Proposition. *Let*

$$\begin{tikzcd} & \rho \arrow[dl, dash] \arrow[dr, dash] & \\ \mu \arrow[dr, dash] & & \lambda \arrow[dl, dash] \\ & \nu & \end{tikzcd}$$

be a fragment of a jeu de taquin growth diagram. (Thus both μ and λ cover ν in the Young lattice, while ρ covers both μ and λ.) Then λ is uniquely determined from ν, μ, and ρ, according to the following rule:

- *if μ is the only shape of its size that contains ν and is contained in ρ, then $\lambda = \mu$;*
- *otherwise there is a unique such shape different from μ, and this is λ.* (A1.136)

In other words, $\mu \neq \lambda$ if and only if the interval $[\nu, \rho]$ in the Young lattice is isomorphic to a product of two 2-element chains.

Proof. Suppose we are given a tableau T of shape λ/μ and a box b such that $\mathrm{jdt}_b(T)$ is well-defined. Encode T as a sequence of shapes, and place these shapes on top of each other – and all together on top of the shape $\mu \setminus \{b\}$, as shown in Figure A1-12. We have to show that repeatedly applying the local rules (A1.136) will produce a tableau (encoded by the lower-right row in Figure A1-12) which is exactly $\mathrm{jdt}_b(T)$. Verification of this reformulation of the definition of jeu de taquin is straightforward, and is left to the reader. □

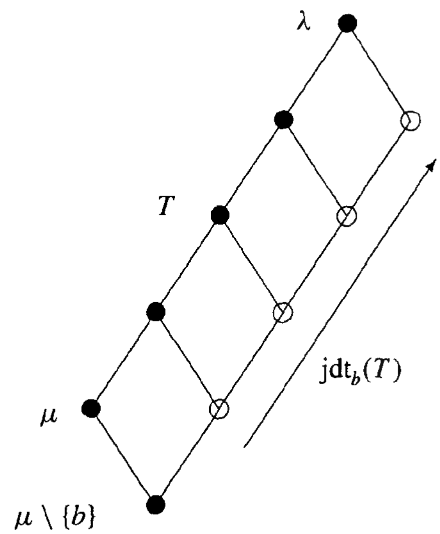


Figure A1-12. Jeu de taquin slide via local transformations.

The growth-diagram interpretation exhibits an important (and not obvious from the original description) *symmetry* of jeu de taquin, which will play a role in the next section. Notice that the rule (A1.136) is symmetric in μ and λ; in other words, λ is computed from ν, μ, and ρ in exactly the same way as μ is computed from ν, λ, and ρ. As a consequence, the "recording" tableau R in Figure A1-11 is equal to $\mathrm{jdt}(S)$, where S is the skew tableau encoded by the upper-right side of the growth diagram (cf. also Figure A1-14).

The Schützenberger Involution

This part of the appendix describes an involution on the set of SYTs of a given shape that is associated with the name of M. P. Schützenberger, and plays an important role in combinatorics, representation theory, and algebraic geometry. The material of this section is not used in the forthcoming proof of the Littlewood–Richardson rule.

A1.2.8 Definition. Let Q be an SYT of shape λ, and let b be its corner box occupied by entry 1. Define

$$\Delta(Q) = \mathrm{jdt}_b(\tilde{Q}) \,,$$

where $\tilde{Q}$ is the skew SYT of shape $\lambda/(1)$ obtained from Q by removing the box b and subsequently decreasing all the remaining entries by 1. The *evacuation tableau* $\mathrm{evac}(Q)$ is by definition the SYT (of shape λ) that is encoded by the sequence of

shapes

$$\emptyset\,,\ \Delta^{n-1}(Q)\,,\ \Delta^{n-2}(Q)\,,\ \ldots\,,\ \Delta^2(Q)\,,\ \Delta(Q)\,,\ Q\,. \qquad \text{(A1.137)}$$

The map $Q \mapsto \operatorname{evac}(Q)$ is called the *Schützenberger involution*. This terminology is justified by the following fact.

A1.2.9 Proposition. *The map $Q \mapsto \operatorname{evac}(Q)$ is an involution.*

Before proving this proposition, let us illustrate Definition A1.2.8 by an example. Take

$$Q = \begin{array}{|c|c|c|}\hline 1&2&4\\ \hline 3&6&7\\ \hline 5\\ \cline{1-1}\end{array}\ . \qquad \text{(A1.138)}$$

Repeatedly applying the operator Δ, we obtain the tableaux

$$\begin{array}{|c|c|c|}\hline 1&3&6\\ \hline 2&5\\ \cline{1-2} 4\\ \cline{1-1}\end{array}\ ,\ \begin{array}{|c|c|c|}\hline 1&2&5\\ \hline 3&4\\ \cline{1-2}\end{array}\ ,\ \begin{array}{|c|c|c|}\hline 1&3&4\\ \hline 2\\ \cline{1-1}\end{array}\ ,\ \begin{array}{|c|c|c|}\hline 1&2&3\\ \hline\end{array}\ ,\ \begin{array}{|c|c|}\hline 1&2\\ \hline\end{array}\ ,\ \begin{array}{|c|}\hline 1\\ \hline\end{array}\ ,\ \emptyset\,. \qquad \text{(A1.139)}$$

The sequence of their shapes (in the reverse order; cf. (A1.137)) encodes the tableau

$$\operatorname{evac}(Q) = \begin{array}{|c|c|c|}\hline 1&2&3\\ \hline 4&5&7\\ \hline 6\\ \cline{1-1}\end{array}\ . \qquad \text{(A1.140)}$$

The reader is encouraged to verify that applying the same procedure to the tableau (A1.140) recovers (A1.138): $\operatorname{evac}(\operatorname{evac}(Q)) = Q$.

Proof. The involution property becomes less mysterious if one reformulates Definition A1.2.8 in terms of growth diagrams. This can be done as follows. The tableaux in (A1.137) can be viewed as sequences of shapes. Let us combine these sequences into a single triangular growth diagram, as shown in Figure A1-13. The rows of this diagram that go in the northeast direction correspond to the tableaux in (A1.137). The whole growth diagram can be reconstructed from its left side (which encodes the original tableau Q) using the local rule (A1.136), together with the fact that all the tableaux in the bottom row are obviously empty. Then

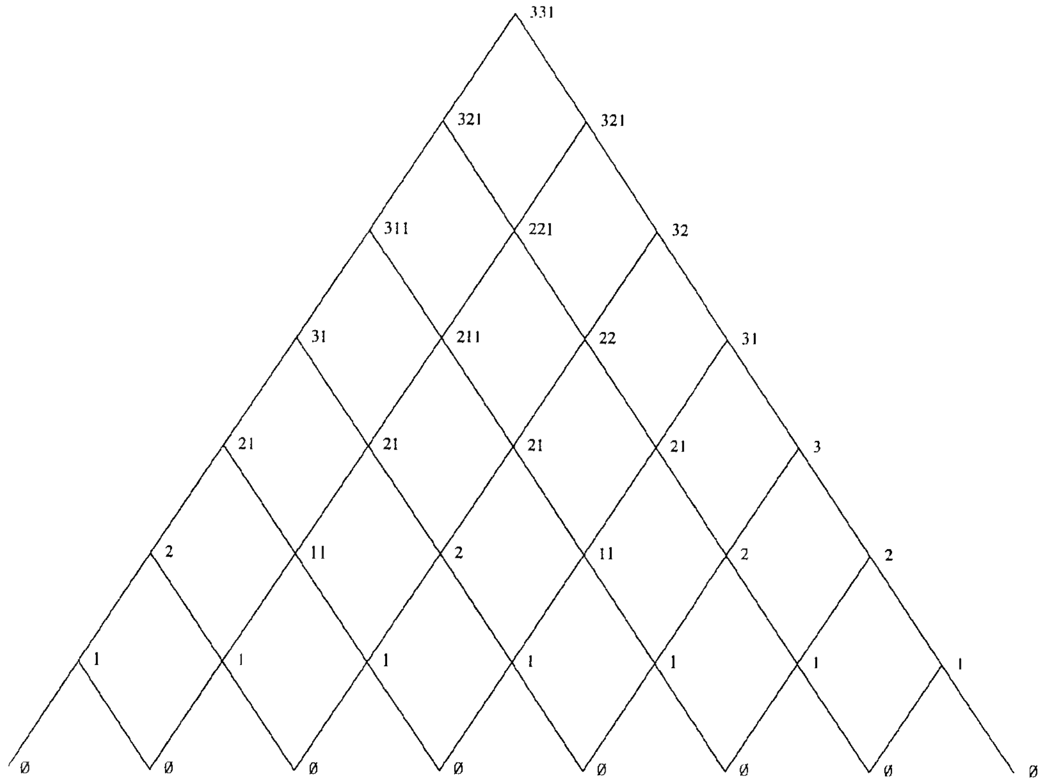


Figure A1-13. The Schützenberger involution.

the right side of the diagram is, by definition, the encoding of evac(Q). Since the rule (A1.136) is symmetric under interchanging λ and μ, applying the same procedure to the tableau evac(Q) would result in the mirror image of the same growth diagram, with its left and right sides interchanged. This proves that $Q \mapsto$ evac(Q) is an involution. $\square$

The following theorem provides a direct interpretation for the Schützenberger involution in terms of the RSK algorithm; it also suggests another proof of Proposition A1.2.9.

For a permutation $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$, let $w^\sharp \in \mathfrak{S}_n$ be given by

$$w^\sharp = n+1-w_n \quad \cdots \quad n+1-w_2 \;\; n+1-w_1.$$

Equivalently, $w^\sharp = w_0 w w_0$, where w_0 denotes the permutation $n \; n-1 \; \cdots \; 2 \; 1$. For example, if $w = 3547126$, then $w^\sharp = 2671435$.

A1.2.10 Theorem. *If* $w \xrightarrow{\text{RSK}} (P, Q)$, *then* $w^\sharp \xrightarrow{\text{RSK}} (\text{evac}(P), \text{evac}(Q))$.

To illustrate, let $w = 3547126$. Then

$$w \xrightarrow{\text{RSK}} (P, Q)\,, \qquad P = \begin{array}{|c|c|c|}\hline 1 & 2 & 6 \\ \hline 3 & 4 & 7 \\ \hline 5 \\ \cline{1-1}\end{array}\,, \qquad Q = \begin{array}{|c|c|c|}\hline 1 & 2 & 4 \\ \hline 3 & 6 & 7 \\ \hline 5 \\ \cline{1-1}\end{array}\,,$$

$$w^\sharp \xrightarrow{\text{RSK}} (P^\sharp, Q^\sharp)\,, \quad P^\sharp = \begin{array}{|c|c|c|}\hline 1 & 3 & 5 \\ \hline 2 & 4 & 7 \\ \hline 6 \\ \cline{1-1}\end{array} = \text{evac}(P)\,, \quad Q^\sharp = \begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 & 7 \\ \hline 6 \\ \cline{1-1}\end{array} = \text{evac}(Q)\,.$$

Proof. The nature of RSK insertion, Knuth equivalence, and jeu de taquin is such that these operations commute with removing the entries that are less than (or larger than) an arbitrary threshold value a. For instance, if we remove all entries of w that are larger than a (thus obtaining a permutation $w_{\le a} \in \mathfrak{S}_a$), then the insertion tableau $P_{\le a}$ of $w_{\le a}$ can be obtained from P by simply removing the boxes containing the entries $a+1, \ldots, n$.

Less trivially, Corollary A1.2.6 implies that the insertion tableau for a permutation $w_{>a}$ is equal to $\text{jdt}(P_{>a})$, where $w_{>a}$ and $P_{>a}$ are obtained by removing the smallest a entries from w and P, respectively, and subtracting a from the remaining entries.

Let $w^\sharp \xrightarrow{\text{RSK}} (P^\sharp, Q^\sharp)$. By Theorem A1.1.1, the shape $\text{sh}(w^\sharp)$ of $P^\sharp$ can be described in terms of the parameters $I_k(w^\sharp)$ that count how many elements can be covered by a union of k increasing subsequences of $w^\sharp$. The argument used above shows that the shape of a partial tableau $P^\sharp_{\le j}$ has a similar description in terms of increasing subsequences of $w^\sharp$ with entries not exceeding j. Note that these subsequences correspond to increasing subsequences of w with entries $> n-j$. Therefore

$$\text{sh}\big(P^\sharp_{\le j}\big) = \text{sh}(w_{>n-j}) = \text{sh}(\text{jdt}(P_{>n-j})) = \text{sh}(\text{evac}(P)_{\le j}),$$

so $P^\sharp = \text{evac}(P)$.

By Theorem 7.13.1, the recording tableau for w coincides with the insertion tableau for w^{-1}. We already proved that as we pass from w to $w^\sharp$, the insertion tableau is replaced by its image under Schützenberger involution. Since $(w^{-1})^\sharp = (w^\sharp)^{-1}$, the same happens to the recording tableau. $\square$

The following corollary of Theorem A1.2.10 is a reformulation of Theorem 7.23.16.

A1.2.11 Corollary. *Let $w = w_1 \cdots w_n \xrightarrow{\text{RSK}} (P, Q)$. Then*

$$ww_0 = w_n \cdots w_1 \xrightarrow{\text{RSK}} (P^t, \text{evac}(Q)^t).$$

Proof. While replacing w by ww_0, we interchange increasing and decreasing subsequences, and each tableau $P_{\leq j}$ gets transposed. Hence the insertion tableau for ww_0 is P^t. As to the recording tableau, we have

$$\begin{array}{rrll} & w_0ww_0 \overset{\text{RSK}}{\longrightarrow} & (\text{evac}(P), \text{evac}(Q)) & \text{(by Theorem A1.2.10)} \\ \Longrightarrow & w_0w^{-1}w_0 \overset{\text{RSK}}{\longrightarrow} & (\text{evac}(Q), \text{evac}(P)) & \text{(by Theorem 7.13.1)} \\ \Longrightarrow & w_0w^{-1} \overset{\text{RSK}}{\longrightarrow} & (\text{evac}(Q)^t, \cdots) & \text{(using what we just proved)} \\ \Longrightarrow & ww_0 \overset{\text{RSK}}{\longrightarrow} & (\cdots, \text{evac}(Q)^t) & \text{(by Theorem 7.13.1),} \end{array}$$

as desired. □

A1.3 The Littlewood–Richardson Rule

The *Littlewood–Richardson coefficients* $c^{\lambda}_{\mu\nu}$ were defined in Section 7.15 (see (7.64)) as the structure constants for the multiplication in the basis of Schur functions:

$$s_\mu s_\nu = \sum_\lambda c^{\lambda}_{\mu\nu} s_\lambda \,, \tag{A1.141}$$

or as coefficients in the expansion of a skew Schur function in this basis:

$$s_{\lambda/\mu} = \sum_\nu c^{\lambda}_{\mu\nu} s_\nu \,. \tag{A1.142}$$

The celebrated Littlewood–Richardson rule is a combinatorial description of the coefficients $c^{\lambda}_{\mu\nu}$. In this section, we prove the rule in two different versions. Three more variations are then stated without proof.

A1.3.1 Theorem (the Littlewood–Richardson rule: jeu de taquin version). *Fix an SYT P of shape ν. The Littlewood–Richardson coefficient $c^{\lambda}_{\mu\nu}$ is equal to the number of SYT of shape λ/μ that are jeu de taquin equivalent to P.*

We will first illustrate Theorem A1.3.1 by an example, and then prove it.

A1.3.2 Example. Let $\lambda = (4, 4, 2, 1)$, $\mu = (2, 1)$, and $\nu = (4, 3, 1)$. Consider the tableau

$$P = \begin{array}{|c|c|c|c|} \hline 1 & 2 & 3 & 4 \\ \hline 5 & 6 & 7 & \\ \cline{1-3} 8 & & & \\ \cline{1-1} \end{array} \tag{A1.143}$$

of shape ν. (According to Theorem A1.3.1, an SYT P of shape ν can be chosen arbitrarily. This special choice of P will later play a role in another version of the

Littlewood–Richardson rule.) There are exactly two SYTs T of shape λ/μ such that $\mathrm{jdt}(T) = P$, namely,

$$\begin{array}{cccc} & & \boxed{3} & \boxed{4} \\ & \boxed{2} & \boxed{6} & \boxed{7} \\ \boxed{1} & \boxed{8} & & \\ \boxed{5} & & & \end{array} \quad \text{and} \quad \begin{array}{cccc} & & \boxed{3} & \boxed{4} \\ & \boxed{2} & \boxed{6} & \boxed{7} \\ \boxed{1} & \boxed{5} & & \\ \boxed{8} & & & \end{array} . \tag{A1.144}$$

Hence $c^{\lambda}_{\mu\nu} = 2$.

Proof of Theorem A1.3.1. We may assume that $|\mu| + |\nu| = |\lambda|$, since otherwise the theorem simply tells that $0 = 0$. Let us then count the number of jeu de taquin growths of the form shown in Figure A1-14. (The shapes λ, μ, and ν, and the tableau P are fixed, while the tableaux R, S, and T are not.)

This number can be found in two different ways, which correspond to reconstructing the growth diagram from its left and right boundary, respectively, using the local rule (A1.136). First, we could count the SYTs T of shape λ/μ such that $P = \mathrm{jdt}(T)$. For each such T, there are f^{μ} possible choices for R. On the other hand, in order to define the values at the right boundary, we only need to pick an

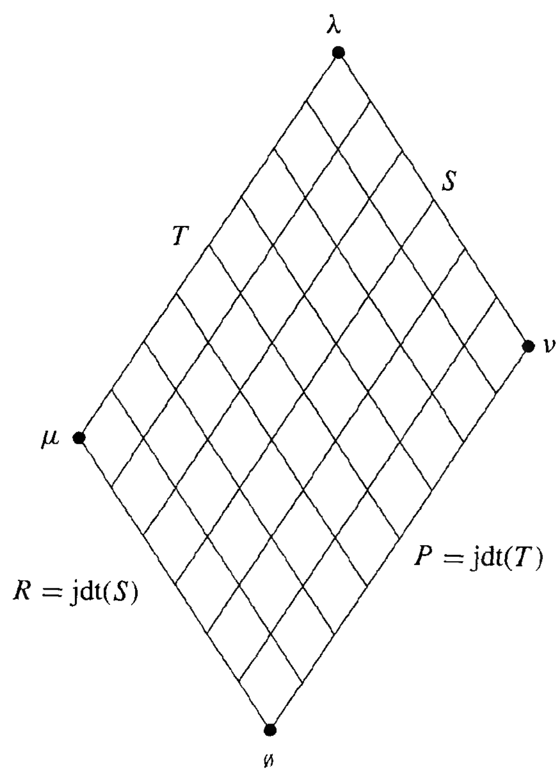


Figure A1-14. Counting jeu de taquin growths.

SYT S of shape λ/ν such that jdt(S) has shape μ. Comparing the two counts, we obtain

$$\begin{aligned} &\#\{\text{SYT } T \text{ of shape } \lambda/\mu \,:\, \mathrm{jdt}(T) = P\} \cdot f^{\mu} \\ &= \#\{\text{SYT } S \text{ of shape } \lambda/\nu \,:\, \mathrm{sh}(\mathrm{jdt}(S)) = \mu\}\,. \end{aligned} \tag{A1.145}$$

This identity implies that the number

$$C^{\lambda}_{\mu\nu} = \#\{\text{SYTs } T \text{ of shape } \lambda/\mu \,:\, \mathrm{jdt}(T) = P\} \tag{A1.146}$$

only depends on the shape ν of P, but not on P itself. (As an aside, notice that the right-hand side of (A1.145) equals $C^{\lambda}_{\nu\mu} \cdot f^{\mu}$, implying that $C^{\lambda}_{\mu\nu} = C^{\lambda}_{\nu\mu}$.)

To prove the theorem, we will need the expansion of a skew Schur function in terms of fundamental quasisymmetric functions that was given in Theorem 7.19.7:

$$s_{\lambda/\mu} = \sum_{T} L_{\mathrm{co}(T)}, \tag{A1.147}$$

where the sum is over all SYTs T of shape λ/μ, and co(T) denotes the composition corresponding to the descent set $D(T)$ of T.

One easily checks that a jeu de taquin slide never changes the relative position of k and $k+1$, implying that the descent set of a skew tableau is invariant under jeu de taquin slides. Therefore the expansion (A1.147) can be rewritten as

$$s_{\lambda/\mu} = \sum_{P} \#\{T \,:\, \mathrm{sh}(T) = \lambda/\mu\,,\ \mathrm{jdt}(T) = P\} \cdot L_{\mathrm{co}(P)},$$

where the sum is over all SYTs P of $n = |\lambda/\mu|$ boxes. In view of (A1.146), this can be further transformed into

$$s_{\lambda/\mu} = \sum_{\nu \vdash n} C^{\lambda}_{\mu\nu} \sum_{P} L_{\mathrm{co}(P)},$$

where the internal sum is over all SYTs P of shape ν. Using (A1.147) for the shape ν, we obtain

$$s_{\lambda/\mu} = \sum_{\nu \vdash n} C^{\lambda}_{\mu\nu}\, s_{\nu}.$$

Comparing this with (A1.142), we conclude that $c^{\lambda}_{\mu\nu} = C^{\lambda}_{\mu\nu}$, which completes the proof of Theorem A1.3.1. □

The Littlewood–Richardson coefficients $c^{\lambda}_{\mu\nu}$ are among the most important families of combinatorial numbers. They appear in the following contexts, among others:

- as coefficients in decompositions of tensor products of irreducible GL_n-modules;
- as coefficients in decompositions of skew Specht modules into irreducibles;
- as coefficients in decompositions of $\mathfrak{S}_n$-representations induced from Young subgroups;
- as intersection numbers in the Schubert calculus on a Grassmannian.

Note that Theorem A1.3.1 readily implies that the Littlewood–Richardson coefficients $c^{\lambda}_{\mu\nu}$ are *nonnegative* integers, a property that is hard to deduce directly from the definitions (A1.141)–(A1.142). Although nonnegativity immediately follows from each of the four interpretations of the $c^{\lambda}_{\mu\nu}$ listed in the previous paragraph, none of these interpretations provides by itself a combinatorial rule that can be used to compute the $c^{\lambda}_{\mu\nu}$. (Note that the third interpretation corresponds to Corollary 7.18.6. Moreover, the first interpretation is discussed in Appendix 2.)

There are many other ways to describe the Littlewood–Richardson coefficients as enumerative combinatorial constants. Once we know that $c^{\lambda}_{\mu\nu}$ is the cardinality of a certain set, then any bijection between this set and another family of combinatorial objects leads to a new description of $c^{\lambda}_{\mu\nu}$. Perhaps the most well-known of such reformulations is the one given in Theorem A1.3.3 below.

Recall from Section 7.10 (see Proposition 7.10.3(d)) that a *lattice permutation* (or Yamanouchi word, or ballot sequence) is a sequence $a_1a_2\cdots a_n$ such that in any initial factor $a_1a_2\cdots a_j$, the number of i's is at least as great as the number of $i+1$'s (for all i). We will also need the notion of a *reverse reading word* of a tableau, which is simply its reading word (cf. Definition A1.1.5) read backwards.

A1.3.3 Theorem (the Littlewood–Richardson rule). *The Littlewood–Richardson coefficient $c^{\lambda}_{\mu\nu}$ is equal to the number of semistandard Young tableaux of shape λ/μ and type ν whose reverse reading word is a lattice permutation.*

A1.3.4 Example. Semistandard Young tableaux with the lattice permutation property described in Theorem A1.3.3 are sometimes called *Littlewood–Richardson tableaux*, or *Littlewood–Richardson fillings* of the shape λ/μ. For the data in Example A1.3.2, there are two such tableaux (thus $c^{\lambda}_{\mu\nu}=2$):

$$\begin{array}{cccc} & & 1 & 1 \\ & 1 & 2 & 2 \\ 1 & 3 & & \\ 2 & & & \end{array} \quad \text{and} \quad \begin{array}{cccc} & & 1 & 1 \\ & 1 & 2 & 2 \\ 1 & 2 & & \\ 3 & & & \end{array}\,. \qquad \text{(A1.148)}$$

The corresponding reverse reading words 11221312 and 11221213 are indeed lattice permutations (of type ν). Note that the Littlewood–Richardson tableaux (A1.148) can be obtained from (A1.144) by replacing the entries 1, 2, 3, 4 by 1, the entries 5, 6, 7 by 2, and the entry 8 by 3.

We are going to deduce Theorem A1.3.3 from Theorem A1.3.1. This will require some preliminary work.

Through the end of this section, a partition $\nu = (\nu_1, \ldots, \nu_k)$ is assumed fixed, and we use the notation

$$N_0 = 0\,, \qquad N_1 = \nu_1\,, \qquad N_2 = \nu_1 + \nu_2\,, \qquad N_3 = \nu_1 + \nu_2 + \nu_3\,, \ldots .$$

Let P_ν denote the particular SYT of shape ν obtained by placing the entries $1, 2, \ldots, n$ in the boxes of ν row by row, beginning with the top row. For instance, if $\nu = (4, 3, 1)$, then P_ν is given by (A1.143). In general, the i-th row of P_ν will be

$$\begin{array}{|c|c|c|c|c|c|}\hline N_{i-1}+1 & N_{i-1}+2 & \cdots & \cdots & \cdots & N_i \\ \hline \end{array}\,, \tag{A1.149}$$

for $i = 1, \ldots, k$.

Let $\mathcal{L}_\nu$ denote the set of all Littlewood–Richardson tableaux of type ν and any shape whatsoever. The following construction will be needed in order to relate the concept of a Littlewood–Richardson tableau to jeu de taquin.

A1.3.5 Definition. Take any SSYT L of type ν (in particular, L could be a Littlewood–Richardson tableau in $\mathcal{L}_\nu$). For any i, the entries of L that are equal to i form a horizontal strip. Replace the 1's in L by $1, \ldots, N_1$, the 2's by $N_1 + 1, \ldots, N_2$, etc., so that the numbers increase left-to-right within each of these horizontal strips. Let us denote the resulting SYT by $\mathrm{st}(L)$ and call it the *standardization* of L. For example, applying this procedure to the Littlewood–Richardson tableaux (A1.148) would give the tableaux in (A1.144).

A1.3.6 Lemma. *A skew SYT T is a standardization of some Littlewood–Richardson tableau of type ν (i.e., $T \in \mathrm{st}(\mathcal{L}_\nu)$) if and only if the following condition holds, for $i = 1, \ldots, k-1$:*

the partial tableaux formed by the entries $N_{i-1}+1, \ldots, N_{i+1}$ of T and P_ν, respectively, are jeu de taquin equivalent. (A1.150)

Proof. First observe that an SYT L is a standardization of some SSYT of type ν (not necessarily a Littlewood–Richardson one) if and only if each of its partial tableaux formed by the entries $N_{i-1}+1, \ldots, N_i$ is jeu de taquin equivalent to the

tableau (A1.149), i.e., to the ith row of P_ν. This condition is obviously satisfied whenever (A1.150) holds. A Littlewood–Richardson tableau L should also satisfy a lattice permutation condition, which is a certain restriction on the partial tableaux formed by the entries of L which are equal to i or $i+1$, for $i = 1, \ldots, k-1$. One easily checks that the standardization map translates this condition into (A1.150). □

A1.3.7 Lemma. *The set* $\mathrm{st}(\mathcal{L}_\nu)$ *of standardizations of Littlewood–Richardson tableaux of type* ν *coincides with the jeu de taquin equivalence class of* P_ν.

Proof. The condition (A1.150) is clearly invariant under jeu de taquin slides; hence the set $\mathrm{st}(\mathcal{L}_\nu)$ is a union of jeu de taquin equivalence classes. By Theorem A1.2.4, it is then enough to show that P_ν is the unique straight-shape tableau in $\mathrm{st}(\mathcal{L}_\nu)$.

Consider the SSYT L of type ν and shape ν obtained by placing i's in row i of ν, for every i. It is straightforward to check that L is the only straight-shape Littlewood–Richardson tableau of type ν. Since $\mathrm{st}(L) = P_\nu$, the lemma follows. □

Proof of Theorem A1.3.3. We need to show that $c^\lambda_{\mu\nu}$ equals the number of Littlewood–Richardson tableaux of shape λ/μ and type ν. This is done as follows:

$$\begin{aligned} c^\lambda_{\mu\nu} &= \#\{\text{SYT } T : \mathrm{sh}(T) = \lambda/\mu,\ T \overset{\text{jdt}}{\sim} P_\nu\} && \text{(by Theorem A1.3.1)} \\ &= \#\{T \in \mathrm{st}(\mathcal{L}_\nu) : \mathrm{sh}(T) = \lambda/\mu\} && \text{(by Lemma A1.3.7)} \\ &= \#\{L \in \mathcal{L}_\nu : \mathrm{sh}(L) = \lambda/\mu\} && \text{(since st is injective on } \mathcal{L}_\nu). \end{aligned}$$

□

Variations of the Littlewood–Richardson Rule

Note that a Littlewood–Richardson coefficient $c^\lambda_{\mu\nu}$ can be defined as a scalar product: $c^\lambda_{\mu\nu} = \langle s_{\lambda/\mu}, s_\nu\rangle$. The following two variations of the Littlewood–Richardson rule, stated without proof, provide combinatorial descriptions of more general *intertwining numbers* $\langle s_\theta, s_\sigma\rangle$, for arbitrary skew shapes θ and σ. (Recall from the discussion after equation (7.64) that $\langle s_\theta, s_\sigma\rangle$ is actually a special case of $\langle s_{\lambda/\mu}, s_\mu\rangle$. However, regarding $\langle s_\theta, s_\sigma\rangle$ in this way obscures the symmetry between θ and σ that appears in Theorems A1.3.8 and A1.3.9 below.)

We will say that a box a is located (weakly) northwest of box b if a occupies a row above b, or the same row as b, and also a column to the left of b, or the same column as b. In a similar fashion, we define what it means for one box to be (weakly) southwest of another.

A1.3.8 Theorem. *For a pair of skew shapes θ and σ, the intertwining number $\langle s_\theta, s_\sigma\rangle$ is equal to the number of bijective maps $f : \theta \to \sigma$ satisfying the following conditions:*

(i) *if box a is located (weakly) northwest of box b, then $f(a)$ is (weakly) southwest of $f(b)$;*
(ii) *if $f(a)$ is located (weakly) northwest of $f(b)$, then a is (weakly) southwest of b.*

(Note that condition (ii) is the same as (i) imposed on the inverse map f^{-1}.)

To illustrate, take the shapes $\theta = (4, 3, 1)$ and $\sigma = (4, 4, 2, 1)/(2, 1)$ (cf. Examples A1.3.2 and A1.3.4). Then there are two bijections $\theta \to \sigma$ satisfying conditions (i)–(ii) of Theorem A1.3.8, which are described by

$$\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & \\ 8 & & & \\ & & & \end{array} \quad \longrightarrow \quad \begin{array}{cccc} & & 8 & 7 \\ & 6 & 3 & 4 \\ 5 & 2 & & \\ 1 & & & \end{array}$$

and

$$\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & \\ 8 & & & \\ & & & \end{array} \quad \longrightarrow \quad \begin{array}{cccc} & & 6 & 7 \\ & 8 & 3 & 4 \\ 5 & 2 & & \\ 1 & & & \end{array}$$

(here each box on the left-hand side is mapped to the box with the same label on the right-hand side). Thus in this case $\langle s_\theta, s_\sigma\rangle = 2$.

A1.3.9 Theorem. *For a pair of skew shapes θ and σ, the intertwining number $\langle s_\theta, s_\sigma\rangle$ is equal to the number of pairs (P, Q) of standard Young tableaux of shapes θ and σ, respectively, such that the reverse reading words of P and Q are permutations inverse to each other.*

(In this theorem, reverse reading words could be replaced by ordinary reading words.)

For $\theta = (4, 3, 1)$ and $\sigma = (4, 4, 2, 1)/(2, 1)$, there are two pairs of tableaux (P, Q) satisfying the conditions of Theorem A1.3.9:

$$P = \begin{array}{cccc} 1 & 2 & 5 & 7 \\ 3 & 4 & 8 & \\ 6 & & & \end{array}, \qquad Q = \begin{array}{cccc} & & 3 & 4 \\ & 2 & 6 & 7 \\ 1 & 8 & & \\ 5 & & & \end{array}$$

and

$$P = \begin{array}{|c|c|c|c|}\hline 1 & 2 & 5 & 7 \\ \hline 3 & 4 & 6 \\ \cline{1-3} 8 \\ \cline{1-1}\end{array}\ , \qquad Q = \begin{array}{cccc} & & \boxed{3} & \boxed{4} \\ & \boxed{2} & \boxed{6} & \boxed{7} \\ \boxed{1} & \boxed{5} & & \\ \boxed{8} & & & \end{array}\ ,$$

with reverse reading words $w = 75218436$, $w^{-1} = 43762815$ and $w = 75216438$, $w^{-1} = 43762518$, respectively.

The last version of the Littlewood–Richardson rule that we are going to discuss exhibits certain symmetries of the coefficients $c^{\lambda}_{\mu\nu}$ that were hidden in the previous versions. Let λ, μ, and ν be partitions with at most r parts satisfying $|\lambda| = |\mu| + |\nu|$. Define the vectors $l = (l_1, \ldots, l_{r-1})$, $m = (m_1, \ldots, m_{r-1})$, and $n = (n_1, \ldots, n_{r-1})$ by

$$\begin{aligned} l_i &= \lambda_{r-i} - \lambda_{r-i+1}\ , \\ m_i &= \mu_i - \mu_{i+1}\ , \\ n_i &= \nu_i - \nu_{i+1}\ . \end{aligned} \tag{A1.151}$$

(It is possible to show that the Littlewood–Richardson coefficient $c^{\lambda}_{\mu\nu}$ is in fact equal to the dimension of the space of SL_r-invariants in the tensor product of three irreducible SL_r-modules naturally associated to l, m, and n.)

The construction below is due to A. Berenstein and A. Zelevinsky, which explains our choice of terminology.

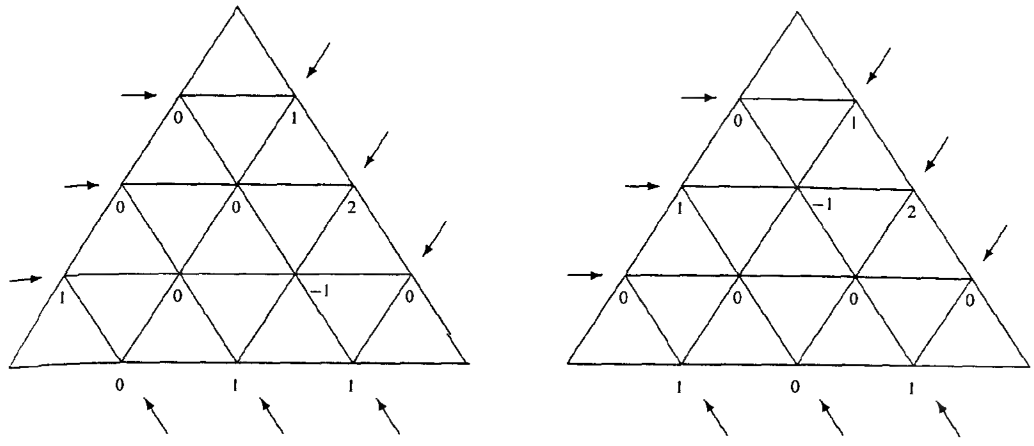
A1.3.10 Definition. Let $l = (l_1, \ldots, l_{r-1})$, $m = (m_1, \ldots, m_{r-1})$, and $n = (n_1, \ldots, n_{r-1})$ be vectors with nonnegative integer components. A *BZ pattern* of type (r, l, m, n) is a collection of integers $(y_{i,j,k})$ indexed by the set

$$\{(i, j, k) \in \mathbb{Z}^3 \ : \ 0 \leq i, j, k < r\ ,\ i + j + k = r\}\ ;$$

and subject to certain linear equations and inequalities to be stated below. It is convenient to view $(y_{i,j,k})$ as a triangular array, as shown in Figure A1-15.

In order to form a BZ pattern, the integers $y_{i,j,k}$ should satisfy the following restrictions. First, the sums of entries along every line of the array that goes in one of the three distinguished directions (excluding the sides of the triangle) are prescribed:

- the sums in the horizontal rows are equal to $l_1, \ldots, l_{r-1}$, top down;
- the sums in the columns going northwest are equal to $m_1, \ldots, m_{r-1}$, left to right;
- the sums in the columns going southwest are equal to $n_1, \ldots, n_{r-1}$, right to left.



$l = (1, 2, 0), m = (1, 1, 0), n = (1, 2, 1)$

Figure A1-15. BZ-patterns for $r = 4, \lambda = (4, 4, 2, 1), \mu = (2, 1, 0, 0), \nu = (4, 3, 1, 0)$.

Second, in each of the $3r - 3$ sums above, the partial sum of several first entries, looking in the direction indicated by an arrow (see Figure A1-15), should be nonnegative. Figure A1-15 shows the two BZ patterns for the data from Examples A1.3.2 and A1.3.4.

A1.3.11 Theorem. *The Littlewood–Richardson coefficient $c^{\lambda}_{\mu\nu}$ is equal to the number of BZ patterns of type (r, l, m, n), where the vectors l, m, and n are defined by* (A1.151).

It is clear from this description that the Littlewood–Richardson coefficient in question is invariant under cyclic permutations of l, m, and n. It is possible to show that $c^{\lambda}_{\mu\nu}$ is in fact *symmetric* as a function of l, m, and n, and is also invariant under simultaneous rearrangement of the entries of each of the vectors l, m, and n in reverse order.

ACKNOWLEDGMENTS. I am grateful to Curtis Greene and Andrei Zelevinsky for a number of valuable suggestions and corrections.

Notes

Theorem A1.1.1 was proved by C. Greene **[6]**, generalizing C. E. Schensted's result **[7.136]**. Corollary A1.1.2 can be extended to arbitrary finite posets, as shown by C. Greene and D. J. Kleitman **[7][8]** (see also **[2]**).

Knuth equivalence and Theorems A1.1.4 and A1.1.6 are due to D. E. Knuth **[7.71]**, who studied this equivalence in a more general setting, with permutations replaced by arbitrary words in the alphabet $\{1, \ldots, n\}$. It is often useful to work in

the *plactic monoid* [**12**], which is the quotient of the free monoid with generators $1, \ldots, n$ under the Knuth equivalence.

Jeu de taquin was invented by M. P. Schützenberger [**17**], as was the involution that bears his name [**7.140**]. Theorem A1.2.4 was proved by M. P. Schützenberger [**17**] and G. Thomas [**18**][**19**], and Theorem A1.2.10 by M. P. Schützenberger [**7.140**]. The growth-diagram interpretation of jeu de taquin was suggested in [**7.32**]. These constructions can be generalized to arbitrary finite posets (cf. [**16**]), although the analogue of Theorem A1.2.4 does not generally hold.

The Littlewood–Richardson rule was discovered by D. E. Littlewood and A. R. Richardson [**7.89**]. First complete proofs were given by M. P. Schützenberger [**17**] and G. Thomas [**18**][**20**]. An incomplete proof published by G. de B. Robinson [**7.131**] and reproduced by D. E. Littlewood [**7.88**, Ch. 6.3, Thm. V] was made precise by I. G. Macdonald [**7.92**, Ch. 1.9][**7.96**, Ch. 1.9] The proof given here is based on a combination of ideas taken from [**17**], [**7.32**], and [**9**].

Theorem A1.3.8 appeared in [**3**], and is a version of a result by A. Zelevinsky [**22**] (cf. also D. E. White [**21**]); the idea goes back to G. D. James and M. H. Peel [**10**]. Theorem A1.3.9 is a result of S. V. Kerov [**11**] and A. M. Garsia and J. B. Remmel [**5**]. Theorem A1.3.11 is due to A. D. Berenstein and A. Zelevinsky [**1**], who also gave reformulations exhibiting other symmetries of the Littlewood–Richardson coefficients.

Other versions of the Littlewood–Richardson rule, along with alternative proofs, can be found in [**3**][**7.35**] [**4**, Ch. 5.2–5.3] [**7.65**, 2.8.13] [**14**] [**15**, Thm. 4.9.4], among other sources. The history of the rule is presented in [**13**, pp. 3–7]. Generalizations and variations of the Littlewood–Richardson rule are numerous, and we do not attempt at reviewing them here.

References

A1.1. A. D. Berenstein and A. V. Zelevinsky, Triple multiplicities for sl$(r+1)$ and the spectrum of the exterior algebra of the adjoint representation, *J. Algebraic Combinatorics* **1** (1992), 7–22.

A1.2. S. V. Fomin, Finite partially ordered sets and Young tableaux, *Soviet Math. Doklady* **19** (1978), 1510–1514.

A1.3. S. Fomin and C. Greene, A Littlewood–Richardson miscellany, *European J. Combinatorics* **14** (1993), 191–212.

A1.4. W. Fulton, *Young Tableaux*, Cambridge University Press, 1997.

A1.5. A. Garsia and J. Remmel, Shuffles of permutations and the Kronecker product, *Graphs and Combinatorics* **1** (1985), 217–263.

A1.6. C. Greene, An extension of Schensted's theorem, *Advances in Math.* **14** (1974), 254–265.

A1.7. C. Greene and D. J. Kleitman, The structure of Sperner k-families, *J. Combinatorial Theory (A)* **20** (1976), 41–68.

A1.8. C. Greene, Some partitions associated with a partially ordered set, *J. Combinatorial Theory (A)* **20** (1976), 69–79.

A1.9. M. D. Haiman, Dual equivalence with applications, including a conjecture of Proctor, *Discrete Math.* **99** (1992), 79–113.

A1.10. G. D. James and M. H. Peel, Specht series for skew representations of symmetric groups, *J. Algebra* **56** (1979), 343–364.

A1.11. S. V. Kerov, The Robinson–Schensted–Knuth correspondence and the Littlewood–Richardson rule, *Russian Math. Surveys* **39** (1984), 165–166.

A1.12. A. Lascoux and M. P. Schützenberger, Le monoïde plaxique, in *Noncommutative Structures in Algebra and Geometric Combinatorics (Naples, 1978), Quad. "Ricerca Sci."* **109**, CNR, Rome, pp. 129–156,.

* A1.13. M. A. A. van Leeuwen, The Littlewood–Richardson rule, theory and implementation, Report MAS-R9709, CWI, Amsterdam, 1997.

A1.14. J. B. Remmel and M. Shimozono, A simple proof of the Littlewood–Richardson rule and applications, *Discrete Math.*, to appear.

A1.15. B. E. Sagan, *The Symmetric Group. Representations, Combinatorial Algorithms, and Symmetric Functions*, Wadsworth & Brooks/Cole, Pacific Grove, California, 1991.

A1.16. M. P. Schützenberger, Promotion des morphismes d'ensembles ordonnés, *Discrete Math.* **2** (1972), 73–94.

A1.17. M. P. Schützenberger, La correspondence de Robinson, in *Combinatoire et Représentation du Groupe Symétrique* (D. Foata, ed.), Lecture Notes in Math. **579**, 1977, pp. 59–135.

A1.18. G. Thomas, *Baxter Algebras and Schur Functions*, Ph.D. thesis, University College of Swansea, 1974.

A1.19. G. Thomas, On a construction of Schützenberger, *Discrete Math.* **17** (1977), 107–118.

A1.20. G. Thomas, On Schensted's construction and the multiplication of Schur functions, *Advances in Math.* **30** (1978), 8–32.

A1.21. D. E. White, Some connections between the Littlewood–Richardson rule and the construction of Schensted, *J. Combinatorial Theory (A)* **30** (1981), 237–247.

A1.22. A. V. Zelevinsky, A generalization of the Littlewood–Richardson rule and the Robinson–Schensted–Knuth correspondence, *J. Algebra* **69** (1981), 82–94.

Chapter 7: Appendix 2

The Characters of GL $(n, \mathbb{C})$

In this appendix we state without proof the fundamental connection between Schur functions and the characters of the general linear group $\mathrm{GL}(n, \mathbb{C})$. By definition, $\mathrm{GL}(n, \mathbb{C})$ is the group of all invertible $n \times n$ matrices with complex entries (under the operation of matrix multiplication). If V is an n-dimensional complex vector space, then after choosing an ordered basis for V we can identify $\mathrm{GL}(n, \mathbb{C})$ with the group $\mathrm{GL}(V)$ of invertible linear transformations $A : V \to V$ (under the operation of composition of linear transformations).

A *linear representation* of $\mathrm{GL}(V)$ is a homomorphism $\varphi : \mathrm{GL}(V) \to \mathrm{GL}(W)$, where W is a complex vector space. From now on we assume that all representations are *finite-dimensional*, i.e., $\dim W < \infty$. We call $\dim W$ the *dimension* of the representation φ, denoted $\dim \varphi$. The representation φ is a *polynomial* (respectively, *rational*) representation if, after choosing ordered bases for V and W, the entries of $\varphi(A)$ are polynomials (respectively, rational functions) in the entries of $A \in \mathrm{GL}(n, \mathbb{C})$. It is clear that the notion of polynomial or rational representation is independent of the choice of ordered bases of V and W, since linear combinations of polynomials (respectively, rational functions) remain polynomials (respectively, rational functions). In general we do not distinguish between representations of $\mathrm{GL}(V)$ and the obvious analogous notion of a representation of $\mathrm{GL}(n, \mathbb{C})$. We say that the representation φ is *homogeneous of degree m* if $\varphi(\alpha A) = \alpha^m \varphi(A)$ for all $\alpha \in \mathbb{C}^* = \mathbb{C} - \{0\}$. If φ is a polynomial (or rational) representation, then this condition is equivalent to saying that each entry of $\varphi(A)$ is a homogeneous polynomial (or rational function) of degree m.

NOTE. Often the dimension of a representation φ is called its *degree*, so do not be confused by our different use of the term "degree."

A2.1 Example. Let $n = 2$, and define $\varphi : \mathrm{GL}(2, \mathbb{C}) \to \mathrm{GL}(3, \mathbb{C})$ by

$$\varphi \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 & 2ab & b^2 \\ ac & ad + bc & bd \\ c^2 & 2cd & d^2 \end{bmatrix}. \qquad \text{(A2.152)}$$

One can directly check that φ is a group homomorphism. Since the entries of the matrix on the right-hand side of (A2.152) are homogeneous polynomials of degree two in a, b, c, d, it follows that φ is a homogeneous polynomial representation of dimension three and degree two.

A2.2 Example. Here are some simple examples of representations illustrating the terms defined above. In all these examples we take $A \in \mathrm{GL}(n, \mathbb{C})$.

- $\varphi(A) = 1 \in \mathbb{C}$ (the *trivial representation*). This is a homogeneous polynomial representation of dimension one and degree zero.
- $\varphi(A) = A$ (the *defining representation*). This is a homogeneous polynomial representation of dimension n and degree one.
- $\varphi(A) = (\det A)^m$, where $m \in \mathbb{Z}$. If $m \geq 0$, then this is a homogeneous polynomial representation of dimension one and degree mn. If $m < 0$ then φ is rational, but not polynomial. The degree remains mn.
- $\varphi(A) = |\det A|^{\sqrt{2}}$. Not a rational representation. It has dimension one and is not homogeneous. (The equation $\varphi(\alpha A) = \alpha^{n\sqrt{2}}\varphi(A)$ only holds when α is a positive real number.)
- $\varphi(A) = A^{-1}$. Not a representation.
- $\varphi(A) = (A^{-1})^t$, where ^{t} denotes transpose. A homogeneous rational (but not polynomial) representation of dimension n and degree -1. To see this, one needs the formula for the entries of the inverse of a matrix mentioned in the proof of Theorem 4.7.2.
- $\varphi(A) = (\det A)^m A$, where $m \in \mathbb{N}$. A homogeneous polynomial representation of dimension n and degree $mn + 1$.
- $\varphi(A) = \bar{A}$, where $^-$ denotes complex conjugation. A nonrational representation of dimension n, and not homogeneous.
- $\varphi(A) = [\sigma(a_{ij})]$, where σ is a field automorphism of $\mathbb{C}$ which is not the identity or complex conjugation (so σ is necessarily discontinuous). This representation (of dimension n) is not only nonrational, but is not continuous.
- $\varphi(A) = \begin{bmatrix} 1 & \log|\det A| \\ 0 & 1 \end{bmatrix}$. A representation of dimension two that isn't homogeneous or rational, though it is continuous.

Note that many of the pathological examples given above disappear when we consider $\mathrm{SL}(n, \mathbb{C}) := \{A \in \mathrm{GL}(n, \mathbb{C}) : \det A = 1\}$ instead of $\mathrm{GL}(n, \mathbb{C})$. We will have more to say about $\mathrm{SL}(n, \mathbb{C})$ later.

Consider the representation φ of Example A2.1. One can check that if the matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

has eigenvalues θ_1 and θ_2, then $\varphi(A)$ has eigenvalues θ_1^2, $\theta_1\theta_2$, and θ_2^2. This computation illustrates the following fundamental result.

A2.3 Proposition. *If φ is a homogeneous rational representation of* $\mathrm{GL}(V)$ *of dimension N and degree m, then there exists a multiset $\mathcal{M}_\varphi$ of N Laurent monomials $x^a = x_1^{a_1} \dots x_n^{a_n}$ of degree m (i.e., $\sum a_i = m$) with the following property. If $A \in \mathrm{GL}(V)$ has eigenvalues $\theta_1, \dots, \theta_n$, then the eigenvalues of $\varphi(A)$ are given by θ^a, for all $x^a \in \mathcal{M}_\varphi$. Moreover, if φ is a polynomial representation, then the Laurent monomials $x^a \in \mathcal{M}_\varphi$ are actual monomials (no negative exponents).*

If φ is a rational representation of $\mathrm{GL}(n, \mathbb{C})$, then define its *character* $\mathrm{char}\,\varphi$ to be the Laurent polynomial

$$\mathrm{char}\,\varphi = (\mathrm{char}\,\varphi)(x) = \sum_{x^a \in \mathcal{M}_\varphi} x^a. \tag{A2.153}$$

Thus if $A \in \mathrm{GL}(n, \mathbb{C})$ has eigenvalues $\theta_1, \dots, \theta_n$, then $(\mathrm{char}\,\varphi)(\theta) = \mathrm{tr}\,\varphi(A)$. Note that an immediate consequence of this definition is the fact that if $\varphi = \varphi_1 \oplus \varphi_2 \oplus \cdots \oplus \varphi_m$ (direct sum of rational representations), then

$$\mathrm{char}\,\varphi = \mathrm{char}\,\varphi_1 + \mathrm{char}\,\varphi_2 + \cdots + \mathrm{char}\,\varphi_m.$$

We are now ready to state the main theorem on rational and polynomial representations of $\mathrm{GL}(n, \mathbb{C})$.

A2.4 Theorem. (I) *Every rational representation $\varphi : \mathrm{GL}(V) \to \mathrm{GL}(W)$ is completely reducible, i.e., every $\mathrm{GL}(V)$-invariant subspace of W has a $\mathrm{GL}(V)$-invariant complement. Hence φ is a direct sum of irreducible representations.*
(II) *Let $\varphi : \mathrm{GL}(V) \to \mathrm{GL}(W)$ and $\varphi' : \mathrm{GL}(V) \to \mathrm{GL}(W')$ be rational representations of $\mathrm{GL}(V)$. Then φ and φ' are equivalent (i.e., there is a bijective linear transformation $\alpha : W \to W'$ such that $\alpha(\varphi(A)(v)) = \varphi'(A)(\alpha(v))$ for all $A \in \mathrm{GL}(V)$ and $v \in W$) if and only if* $\mathrm{char}\,\varphi = \mathrm{char}\,\varphi'$.
(III) *Irreducible rational representations φ of* $\mathrm{GL}(V)$ *are homogeneous, and* $\mathrm{char}\,\varphi$ *is a* symmetric *(homogeneous) Laurent polynomial in $x_1, \dots, x_n$. Hence* $\mathrm{char}\,\psi \in \Lambda_n^m$ *if ψ is a homogeneous polynomial representation of degree m.*
(IV) (The main result.) *The irreducible polynomial representations φ^λ of* $\mathrm{GL}(V)$ *can be indexed by partitions λ of length at most n so that*

$$\mathrm{char}\,\varphi^\lambda = s_\lambda(x_1, \dots, x_n).$$

(V) *Every irreducible rational representation φ of* $\mathrm{GL}(V)$ *is of the form $\varphi(A) = (\det A)^m \varphi'(A)$ for some $m \in \mathbb{Z}$ and some irreducible polynomial representation φ' of* $\mathrm{GL}(V)$. *The corresponding characters are hence related by*

$$\mathrm{char}\,\varphi = (x_1 \cdots x_n)^m\, \mathrm{char}\,\varphi'.$$

It follows from the above theorem that if φ is a polynomial representation of

$GL(V)$, then the multiplicity of the irreducible character φ^λ in φ is given by

$$\langle \varphi, \varphi^\lambda \rangle = \langle \operatorname{char} \varphi, s_\lambda \rangle.$$

Let us consider some simple examples.

- If $\varphi(A) = 1$ (the trivial representation), then $\operatorname{char}\varphi = s_\emptyset = 1$. The representation is irreducible. (The main theorem is hardly needed for irreducibility – any linear representation of dimension one of any group is irreducible.)
- If $\varphi(A) = A$ (the defining representation), then $\operatorname{char} \varphi = x_1 + \cdots + x_n = s_1$ (understood to be in the variables $x_1, \ldots, x_n$). The representation is irreducible.
- If $\varphi(A) = (\det A)^m$ for $m \in \mathbb{Z}$, then $\operatorname{char} \varphi = (x_1 \cdots x_n)^m$. If $m \geq 0$, then $\operatorname{char}\varphi = s_\lambda$, where $\lambda = \langle m^n \rangle$. The representation is irreducible for any m.
- If $\varphi(A) = (A^{-1})^t$, then $\operatorname{char}\varphi = (x_1 \cdots x_n)^{-1} = (x_1 \cdots x_n)^{-1} s_\emptyset$. The representation is irreducible.

For a somewhat more substantial example, let $\operatorname{End}(V)$ denote the set of all linear transformations $X : V \to V$. Consider the action of $GL(V)$ on $\operatorname{End}(V)$ given by $A \cdot X = AXA^{-1}$. This is called the *adjoint representation* of $GL(V)$, denoted ad. Note that $\dim \operatorname{ad} = \dim \operatorname{End}(V) = n^2$. To compute char(ad), first choose an ordered basis for V, so we can identify $\operatorname{End}(V)$ with $\operatorname{Mat}(n)$, the ring of all $n \times n$ complex matrices. Let $A = \operatorname{diag}(\theta_1, \ldots, \theta_n)$, the diagonal matrix with diagonal entries $\theta_1, \ldots, \theta_n$. Let E_{ij} be the matrix in $\operatorname{Mat}(n)$ with a 1 in the (i, j)-position, and 0's elsewhere. Observe that

$$AE_{ij}A^{-1} = \theta_i \theta_j^{-1} E_{ij}.$$

Hence the E_{ij}'s are eigenvectors for $\operatorname{ad}(A)$, with eigenvalues $\theta_i\theta_j^{-1}$. We have found n^2 linearly independent eigenvectors, so

$$\operatorname{tr} \operatorname{ad}(A) = \sum_{i,j} \theta_i \theta_j^{-1},$$

where i, j range from 1 to n. It follows that

$$\begin{aligned}
\operatorname{char}(\operatorname{ad}) &= \sum_{i,j} x_i x_j^{-1} \\
&= 1 + (x_1 \cdots x_n)^{-1} \left[(n-1)(x_1 \cdots x_n) + \sum_{i \neq j} \frac{x_i}{x_j} (x_1 \cdots x_n) \right] \\
&= s_\emptyset + (x_1 \cdots x_n)^{-1} s_{21^{n-2}}. \qquad \text{(A2.154)}
\end{aligned}$$

It follows that ad has two irreducible components, one being the trivial representation (with character $s_\emptyset$). In other words, the space $\operatorname{Mat}(n)$ contains a one-

dimensional subspace invariant under the action of GL(n, $\mathbb{C}$). This space consists of the scalar multiples of the identity matrix. The GL(n, $\mathbb{C}$)-invariant irreducible subspace complementary to these diagonal matrices consists of the matrices of trace 0.

The classical definition of Schur functions (Theorem 7.15.1) may be regarded as giving a formula for char φ^λ as a quotient $a_{\lambda+\delta}/a_\delta$ of two determinants. Readers familiar with the representation theory of semisimple Lie algebras or Lie groups will recognize this formula as the *Weyl character formula* for the group GL(n, $\mathbb{C}$). The factorization $a_\delta = \prod_{i<j}(x_i - x_j)$ of the denominator is just the *Weyl denominator formula* for GL(n, $\mathbb{C}$). Now note that for any polynomial (or rational) representation φ of GL(n, $\mathbb{C}$), we have by the definition (A2.153) of char φ that

$$\dim\varphi = (\mathrm{char}\,\varphi)(1^n). \tag{A2.155}$$

Thus in particular $s_\lambda(1^n) = \dim \mathrm{char}\,\varphi^\lambda$. The formula for $s_\lambda(1^n)$ obtained by substituting $q = 1$ in equation (7.105) is equivalent to the *Weyl dimension formula* for GL(n, $\mathbb{C}$). Corollary 7.21.4 is an alternative form of this formula.

Idea of Proof of Theorem A2.4. Although we will not prove Theorem A2.4 here, let us say a few words about the structure of the proof. Let V be an n-dimensional complex vector space. Then GL(V) acts diagonally on the k-th tensor power $V^{\otimes k}$, i.e,

$$A \cdot (v_1 \otimes \cdots \otimes v_k) = (A \cdot v_1) \otimes \cdots \otimes (A \cdot v_k), \tag{A2.156}$$

and the symmetric group $\mathfrak{S}_k$ acts on $V^{\otimes k}$ by permuting tensor coordinates, i.e.,

$$w \cdot (v_1 \otimes \cdots \otimes v_k) = v_{w^{-1}(1)} \otimes \cdots \otimes v_{w^{-1}(k)}. \tag{A2.157}$$

The actions of GL(V) and $\mathfrak{S}_k$ commute, so we have an action of $\mathfrak{S}_k \times \mathrm{GL}(V)$ on $V^{\otimes k}$. A crucial fact is that the actions of GL(V) and $\mathfrak{S}_k$ *centralize* each other, i.e., the (invertible) linear transformations $V^{\otimes k} \to V^{\otimes k}$ that commute with the $\mathfrak{S}_k$ action are just those given by (A2.156), while conversely the linear transformations that commute with the GL(V) action are those generated (as a $\mathbb{C}$-algebra) by (A2.157). From this it can be shown that $V^{\otimes k}$ decomposes into irreducible $\mathfrak{S}_k \times \mathrm{GL}(V)$-modules as follows:

$$V^{\otimes k} = \coprod_\lambda (M^\lambda \otimes F^\lambda), \tag{A2.158}$$

where $\coprod$ denotes direct sum. Here the M^λ's are nonisomorphic irreducible $\mathfrak{S}_k$ modules, the F^λ's are nonisomorphic irreducible GL(V) modules, and λ ranges over some index set. We know (Theorem 7.18.5) that the irreducible representations of $\mathfrak{S}_k$ are indexed by partitions λ of k, so we choose the indexing so that M^λ is the irreducible $\mathfrak{S}_k$-module corresponding to $\lambda \vdash k$ via Theorem 7.18.5. Thus we have constructed irreducible (or possibly 0) GL(V)-modules F^λ. These modules

afford polynomial representations φ^λ, and the nonzero ones are inequivalent. (The argument below shows that $F^\lambda \neq 0$ if and only if $\ell(\lambda) \leq n$.)

Next we compute the character of φ^λ. Let $w \times A$ be an element of $\mathfrak{S}_k \times \mathrm{GL}(V)$, and let $\mathrm{tr}(w \times A)$ denote the trace of $w \times A$ acting on $V^{\otimes k}$. Then by equation (A2.158) we have

$$\mathrm{tr}(w \times A) = \sum_\lambda \chi^\lambda(w) \cdot \mathrm{tr}(\varphi^\lambda(A)).$$

Let A have eigenvalues $\theta = (\theta_1, \ldots, \theta_n)$. A straightforward computation shows that $\mathrm{tr}(w \times A) = p_{\rho(w)}(\theta)$, so

$$p_{\rho(w)}(\theta) = \sum_\lambda \chi^\lambda(w)(\mathrm{char}\,\varphi^\lambda)(\theta).$$

But we know (Corollary 7.17.4) that

$$p_{\rho(w)} = \sum_\lambda \chi^\lambda(w) s_\lambda.$$

Since the χ^λ's are linearly independent, we conclude $\mathrm{char}\,\varphi^\lambda = s_\lambda$. A separate argument shows that there are no other irreducible polynomial characters, and the (sketched) proof is complete. $\square$

The *special linear group* $\mathrm{SL}(n, \mathbb{C})$ is defined to be the subgroup of $\mathrm{GL}(n, \mathbb{C})$ consisting of the matrices in $\mathrm{GL}(n, \mathbb{C})$ of determinant one. It is sometimes more convenient to work with $\mathrm{SL}(n, \mathbb{C})$ rather than $\mathrm{GL}(n, \mathbb{C})$, so we will discuss the basics of the representation theory of $\mathrm{SL}(n, \mathbb{C})$. The main result is the following.

A2.5 Theorem. *Let $\lambda \in \mathrm{Par}$ with $\ell(\lambda) \leq n$. Then the restriction of φ^λ to $\mathrm{SL}(n, \mathbb{C})$ remains irreducible. The representations φ^λ for $\ell(\lambda) \leq n-1$ are all inequivalent. If $\ell(\lambda) = n$, then*

$$\varphi^\lambda = \varphi^{(\lambda_1 - \lambda_n, \lambda_2 - \lambda_n, \ldots, 0)},$$

as representations of $\mathrm{SL}(n, \mathbb{C})$. Every polynomial (or rational) representation of $\mathrm{SL}(n, \mathbb{C})$ is a direct sum of irreducible representations φ^λ for $\ell(\lambda) \leq n-1$.

If $\theta_1, \ldots, \theta_n$ are the eigenvalues of $A \in \mathrm{SL}(n, \mathbb{C})$, then $\theta_1 \cdots \theta_n = 1$. Hence it is natural to define the character $\mathrm{char}\,\varphi$ of a polynomial representation φ of $\mathrm{SL}(n, \mathbb{C})$ as lying in the ring $\Xi_n = \Lambda_n/(x_1 \cdots x_n - 1)$, the ring of symmetric functions in the variables $x_1, \ldots, x_n$, modulo the relation $x_1 \cdots x_n = 1$. A $\mathbb{C}$-basis for this ring consists of Schur functions s_λ with $\ell(\lambda) \leq n-1$, and two polynomial representations of $\mathrm{SL}(n, \mathbb{C})$ are equivalent if and only if they have the same character (regarded as lying in Ξ_n).

As an example of a computation of an $\mathrm{SL}(n, \mathbb{C})$ character, define the *adjoint representation* of $\mathrm{SL}(n, \mathbb{C})$ to be the action of $\mathrm{SL}(n, \mathbb{C})$ on the set $\mathrm{End}_0(V)$ of

linear transformations $X : V \to V$ of trace 0 given by $A \cdot X = AXA^{-1}$. In other words, we restrict the adjoint representation of $\mathrm{GL}(n, \mathbb{C})$ to the subgroup $\mathrm{SL}(n, \mathbb{C})$ and its action to the subspace of $\mathrm{End}(V)$ consisting of linear transformations of trace 0 (the Lie algebra $\mathfrak{sl}(V)$). From equation (A2.154) we saw that $\mathrm{End}_0(V)$ was an irreducible subspace for the adjoint action of $\mathrm{GL}(n, \mathbb{C})$, with character $(x_1 \cdots x_n)^{-1} s_{21^{n-2}} \in \Lambda_n$. Hence the adjoint action of $\mathrm{SL}(n, \mathbb{C})$ is irreducible, with character $s_{21^{n-2}} \in \Xi_n$.

We conclude this appendix by discussing the connections between representation theory and two operations on symmetric functions. The first operation is the usual product fg. Let $\varphi : \mathrm{GL}(V) \to \mathrm{GL}(W)$ and $\varphi' : \mathrm{GL}(V) \to \mathrm{GL}(W')$ be two polynomial representations of $\mathrm{GL}(V)$. The tensor product representation

$$\varphi \otimes \varphi' : \mathrm{GL}(V) \to \mathrm{GL}(W \otimes W')$$

is defined by $A \cdot (w \otimes w') = (A \cdot w) \otimes (A \cdot w')$ (and extended to all of $W \otimes W'$ by bilinearity). If $B : W \to W$ and $B' : W' \to W'$ are linear transformations with eigenvalues $\theta_1, \ldots, \theta_N$ and $\theta'_1, \ldots, \theta'_{N'}$, respectively, then the eigenvalues of $B \otimes B'$ are the numbers $\theta_i \theta'_j$. It follows that

$$\mathrm{char}(\varphi \otimes \varphi') = \mathrm{char}\, \varphi \cdot \mathrm{char}\, \varphi'.$$

In particular if λ, μ, ν are partitions of length at most $n = \dim V$, then the multiplicity of φ^λ in $\varphi^\mu \otimes \varphi^\nu$ is just the Littlewood–Richardson coefficient

$$\langle s_\lambda, s_\mu s_\nu \rangle = c^\lambda_{\mu\nu}.$$

Hence the Littlewood–Richardson coefficients have a simple interpretation involving the representation theory of $\mathrm{GL}(n, \mathbb{C})$, showing in particular that they are nonnegative. Thus we have seen three fundamental ways to show that $c^\lambda_{\mu\nu} \geq 0$: (a) combinatorially, *via* the Littlewood–Richardson rule (Theorem A1.3.1), (b) algebraically, using the representation theory of the symmetric group (see Corollary 7.18.6), and (c) algebraically, using the representation theory of the general linear group (as done just above). The two methods (b) and (c) are in a sense "dual" to each other, the duality arising from the pairing (A2.158) of irreducible representations of $\mathfrak{S}_k$ and $\mathrm{GL}(V)$.

The second operation on symmetric functions we are considering here would appear at first sight rather unmotivated without understanding the connection with representation theory. Suppose that we have polynomial representations $\varphi : \mathrm{GL}(V) \to \mathrm{GL}(W)$ and $\psi : \mathrm{GL}(W) \to \mathrm{GL}(Y)$. Then the composition $\psi\varphi : \mathrm{GL}(V) \to \mathrm{GL}(Y)$ defines a polynomial representation of $\mathrm{GL}(V)$. We want to compute its character. Suppose that $A \in \mathrm{GL}(V)$ has eigenvalues $\theta_1, \ldots, \theta_n$. Then by Proposition A2.3 the eigenvalues of $\varphi(A)$ are the monomials θ^a for $x^a \in \mathcal{M}_\varphi$. Similarly, if B has eigenvalues $\zeta_1, \ldots, \zeta_N$, then the eigenvalues of $\psi(B)$ are the monomials ζ^b for $x^b \in \mathcal{M}_\psi$. Hence if we denote the monomials θ^a by $\theta^{a^1}, \ldots, \theta^{a^N}$ (in some

order), then the eigenvalues of $\psi\varphi(A)$ are just the monomials

$$x^b|_{x_i=\theta^{a^i}}, \quad x^b \in \mathcal{M}_\psi.$$

Thus if $f = \sum_{i=1}^N \theta^{a^i} = \operatorname{char}\varphi$ and $g = \operatorname{char}\psi$, then we get

$$\operatorname{char}(\psi\varphi) = g(\theta^{a^1}, \dots, \theta^{a^N}). \tag{A2.159}$$

This formula leads us to the following definition.

A2.6 Definition. Suppose that the symmetric function $f \in \Lambda$ is a sum of monomials, say, $f = \sum_{i\geq 1} x^{a^i}$. Given $g \in \Lambda$, define the *plethysm* $g[f]$ (sometimes denoted $f \circ g$) by

$$g[f] = g(x^{a^1}, x^{a^2}, \dots).$$

(For the etymology of the term "plethysm," see the Notes.)

For instance, since $s_1 = x_1 + x_2 + \cdots$, we have $g[s_1] = g(x_1, x_2, \dots) = g$. More generally, from $p_n = x_1^n + x_2^n + \cdots$ we have

$$f[p_n] = f(x_1^n, x_2^n, \dots) = \sum_{i\geq 1} x^{a^i n} = p_n[f]. \tag{A2.160}$$

Clearly by definition of plethysm we have

$$(af + bg)[h] = af[h] + bg[h], \quad a, b \in \mathbb{Q} \tag{A2.161}$$

$$(fg)[h] = f[h] \cdot g[h]. \tag{A2.162}$$

We can use equation (A2.160) to define $p_n[f]$ for *any* $f \in \Lambda$, and then equations (A2.161) and (A2.162) allow us to define $g[f]$ for any $f, g \in \Lambda$. Specifically, if $g = \sum_\lambda c_\lambda p_\lambda$, then

$$g[f] = \sum_\lambda c_\lambda \prod_{i=1}^{\ell(\lambda)} f(x_1^{\lambda_i}, x_2^{\lambda_i}, \dots).$$

For instance, using Proposition 7.7.6 and equation (7.19), we have

$$\begin{aligned} h_n[-s_1] &= \sum_{\lambda\vdash n} z_\lambda^{-1} p_\lambda[-s_1] \\ &= \sum_{\lambda\vdash n} z_\lambda^{-1} \prod_{i=1}^{\ell(\lambda)} (-s_1(x_1^{\lambda_i})) \\ &= \sum_{\lambda\vdash n} z_\lambda^{-1} (-1)^{\ell(\lambda)} p_\lambda \\ &= (-1)^n e_n. \end{aligned}$$

Hence from equations (A2.161) and (A2.162) it follows that

$$f[-s_1] = (-1)^n \omega(f) \qquad \text{for any } f \in \Lambda^n. \tag{A2.163}$$

Certain symmetric function identities can often be recast in the language of plethysm. As just one example, Corollary 7.13.8 states that

$$\frac{1}{\prod_i (1 - x_i) \cdot \prod_{i<j} (1 - x_i x_j)} = \sum_\lambda s_\lambda(x).$$

We leave for the reader to see that this identity is equivalent to the plethystic formula

$$\left(\sum_{n \geq 0} h_n \right) [e_1 + e_2] = \sum_\lambda s_\lambda. \tag{A2.164}$$

Returning to the representation-theoretic significance of plethysm, when we compare equation (A2.159) with the definition of plethysm (Definition A2.4), we see that

$$\mathrm{char}(\psi\varphi) = (\mathrm{char}\, \psi)[\mathrm{char}\, \varphi].$$

Here it is understood that the variables are restricted to $x_1, \ldots, x_n$ (where φ is a representation of $\mathrm{GL}(n, \mathbb{C})$). Many symmetric function identities thereby acquire a representation-theoretic significance. For instance, let $H(x) = h_0 + h_1 + \cdots$. Readers with a sufficient algebraic background will recognize $H(x_1, \ldots, x_n)$ as the character of $\mathrm{GL}(V)$, where $\dim V = n$, acting on the symmetric algebra $S(V)$. Similarly, $(e_1 + e_2)(x_1, \ldots, x_n)$ is the character of $\mathrm{GL}(V)$ acting on $V \oplus \Lambda^2(V)$, where Λ^2 denotes the second exterior power. Hence equation (A2.164) is equivalent to the assertion that the action of $\mathrm{GL}(V)$ on $S(V \oplus \Lambda^2(V))$ contains every irreducible polynomial representation of $\mathrm{GL}(V)$ exactly once (and contains no non-polynomial representation).

An important property of plethysm is given by the following result.

A2.7 Theorem. *Let f and g be $\mathbb{N}$-linear combinations of Schur functions. Then the plethysm $g[f]$ is also an $\mathbb{N}$-linear combination of Schur functions.*

Proof. Since f is an $\mathbb{N}$-linear combination of Schur functions, for any $m \in \mathbb{P}$ we have that $f(x_1, \ldots, x_m)$ is the character of a polynomial representation φ : $\mathrm{GL}(m, \mathbb{C}) \to \mathrm{GL}(n, \mathbb{C})$ for n sufficiently large. Similarly, $g(x_1, \ldots, x_n)$ is the character of a polynomial representation ψ : $\mathrm{GL}(n, \mathbb{C}) \to \mathrm{GL}(Y)$. Hence the plethysm $g[f](x_1, \ldots, x_n)$ is the character of the composition $\psi\varphi$ and so is an $\mathbb{N}$-linear combination of Schur functions. Now let $n \to \infty$. $\square$

No proof is known of Theorem A2.7 that doesn't use representation theory. In particular, a combinatorial rule for expanding the plethysm $s_n[s_m]$, analogous to the Littlewood–Richardson rule, is not known. Finding such a rule remains one of the outstanding open problems in the theory of symmetric functions.

Just as we defined in Section 7.18 an "induction product" of characters of symmetric groups to correspond to ordinary product of symmetric functions, we can ask whether there is some kind of product of $\mathfrak{S}_n$ characters that corresponds to plethysm. We will briefly sketch the answer to this question, assuming some knowledge of group theory. Suppose that $n = km$. We may regard the group $\mathfrak{S}_k^m$ as a (Young) subgroup of $\mathfrak{S}_n$ in a natural way. Let $N = N(\mathfrak{S}_k^m)$ denote the normalizer of $\mathfrak{S}_k^m$ in $\mathfrak{S}_n$. Thus N is isomorphic to the *wreath product* $\mathfrak{S}_k \,\S\, \mathfrak{S}_m$ (also denoted $\mathfrak{S}_k \wr \mathfrak{S}_m$, $\mathfrak{S}_k \sim \mathfrak{S}_m$, $\mathfrak{S}_k \operatorname{wr} \mathfrak{S}_m$, and $\mathfrak{S}_m[\mathfrak{S}_k]$). Given representations $\sigma : \mathfrak{S}_k \to \mathrm{GL}(V)$ and $\rho : \mathfrak{S}_m \to \mathrm{GL}(W)$, there is a natural (functorial) way to define a representation $\sigma \,\S\, \rho : N \to \mathrm{GL}(V^{\otimes m} \otimes W)$, as follows: we can regard in an obvious way an element of N as a pair (f, v), where $f : [m] \to \mathfrak{S}_k$ and $v \in \mathfrak{S}_m$. Then define

$$(f, v) \cdot (x_1 \otimes \cdots \otimes x_m \otimes y) = \left(f(1)x_{v^{-1}(1)}\right) \otimes \cdots \otimes \left(f(m)x_{v^{-1}(m)}\right) \otimes (v \cdot y).$$

If χ^α denotes the character of a representation α, then set $\chi^\sigma \,\S\, \chi^\rho = \chi^{\sigma \S \rho}$. We can now state the main result connecting plethysm with representations of $\mathfrak{S}_n$.

A2.8 Theorem. *Let χ be a character of $\mathfrak{S}_k$ and θ a character of $\mathfrak{S}_m$. Then*

$$\operatorname{ch} \operatorname{ind}_{N(\mathfrak{S}_k^m)}^{\mathfrak{S}_{km}} (\chi \,\S\, \theta) = (\operatorname{ch} \theta)[\operatorname{ch} \chi].$$

A2.9 Example. A *1-factor* (or *complete matching*) on $[2n]$ is a graph with the vertex set $[2n]$ and with n vertex-disjoint edges. Let O_n denote the set of all 1-factors on $[2n]$. The symmetric group $\mathfrak{S}_{2n}$ acts on O_n by permuting vertices. How does the character ψ^n of this action decompose into irreducibles? The action of $\mathfrak{S}_{2n}$ on O_n is transitive, and the stabilizer of the 1-factor with edges $\{1, 2\}$, $\{3, 4\}, \ldots, \{2n-1, 2n\}$ is precisely the subgroup $N(\mathfrak{S}_2^n)$. Hence

$$\operatorname{ch} \psi^n = \operatorname{ch} 1_{N(\mathfrak{S}_2^n)}^{\mathfrak{S}_{2n}} = \left(\operatorname{ch} 1_{\mathfrak{S}_n}\right)\left[\operatorname{ch} 1_{\mathfrak{S}_2}\right] = h_n[h_2].$$

Now

$$\sum_{n \geq 0} h_n[h_2] = \prod_{i \leq j} (1 - x_i x_j)^{-1}.$$

Putting $q = 0$ in equation (7.202) shows that

$$\prod_{i \leq j} (1 - x_i x_j)^{-1} = \sum_{\mu \vdash n} s_{2\mu}.$$

Hence we get

$$\langle \psi^n, \chi^\lambda \rangle = \begin{cases} 1 & \text{if } \lambda = 2\mu \text{ for some } \mu \vdash n \\ 0 & \text{otherwise.} \end{cases}$$

Exercises

7.1. [1] True or false? The Ferrers diagram of the partition (4, 4, 3) is given by

$$\begin{array}{cccc} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array}$$

7.2. Let Par(n) denote the set of all partitions of n with the dominance ordering.

a. [2] Show that Par(n) is a lattice.

b. [2+] Show that Par(n) is self-dual.

c. [2+] Find the smallest value of n for which Par(n) is not graded.

d. [2+] Show that the maximum number of elements covered by an element of Par(n) is $\lfloor \frac{1}{2}(\sqrt{1+8n} - 3) \rfloor$.

e. [2+] Show that the shortest maximal chain in Par(n) has length $2n - 4$ for $n \geq 3$.

f. [3−] Show that the longest maximal chain in Par(n) has length

$$\tfrac{1}{3}m(m^2 + 3r - 1) \sim \tfrac{1}{3}(2n)^{3/2},$$

where $n = \binom{m+1}{2} + r,\ 0 \leq r \leq m$.

7.3. [2+] Expand the power series $\prod_{i \geq 1}(1 + x_i + x_i^2)$ in terms of the elementary symmetric functions.

7.4. [2+] Show that

$$h_r(x_1, \ldots, x_n) = \sum_{k=1}^{n} x_k^{n-1+r} \prod_{i \neq k} (x_k - x_i)^{-1}.$$

7.5. [2+] Prove the identity

$$\left(1 - \sum_{n \geq 1} p_n t^n\right)^{-1} = \frac{\sum_{n \geq 0} h_n t^n}{1 - \sum_{n \geq 1}(n-1)h_n t^n}. \tag{7.165}$$

7.6. [2+] Let $w \in \mathfrak{S}_n$ have cycle type λ. Give a direct bijective proof of Proposition 7.7.3, i.e., the number of elements $v \in \mathfrak{S}_n$ commuting with w is equal to $z_\lambda = 1^{m_1} m_1! \, 2^{m_2} m_2! \cdots$, where $m_i = m_i(\lambda)$.

7.7. [2+] Let Ω^n denote the subspace of Λ^n consisting of all $f \in \Lambda^n$ satisfying

$$f(x_1, -x_1, x_3, x_4, \ldots) = f(x_3, x_4, \ldots).$$

For instance, $m_1 = x_1 + x_2 + \cdots \in \Omega^1$. Find a "simple" basis for Ω^n. Express the dimension of Ω^n in terms of the number of partitions of n with a suitable restriction.

7.8. [2+] Let $f \in \Lambda^n$, and for any $g \in \Lambda^n$ define $g_k \in \Lambda^{nk}$ by

$$g_k(x_1, x_2, \ldots) = g\left(x_1^k, x_2^k, \ldots\right).$$

Show that

$$\omega f_k = (-1)^{n(k-1)}(\omega f)_k.$$

7.9. [2+] Let λ be a partition of n of length ℓ. Define the *forgotten symmetric function* f_λ by

$$f_\lambda = \varepsilon_\lambda\, \omega(m_\lambda),$$

where $\varepsilon_\lambda = (-1)^{n-\ell}$ as usual. (Sometimes f_λ is defined just as $\omega(m_\lambda)$.) Let $f_\lambda = \sum_\mu a_{\lambda\mu} m_\mu$. Show that $a_{\lambda\mu}$ is equal to the number of distinct permutations $(\alpha_1, \alpha_2, \ldots, \alpha_\ell)$ of the sequence $(\lambda_1, \lambda_2, \ldots, \lambda_\ell)$ such that

$$\{\alpha_1 + \alpha_2 + \cdots + \alpha_i : 1 \leq i \leq \ell\} \supseteq \{\mu_1 + \mu_2 + \cdots + \mu_j : 1 \leq j \leq \ell(\mu)\}.$$

For instance, if $\lambda = (3, 2, 1, 1)$ and $\mu = (5, 2)$, then $a_{\lambda\mu} = 5$, corresponding to $(3, 2, 1, 1)$, $(2, 3, 1, 1)$, $(1, 1, 3, 2)$, $(1, 3, 1, 2)$, and $(3, 1, 1, 2)$.

7.10. [3−] Let $\lambda \in \mathrm{Par}$, and define the symmetric power series

$$A_\lambda(x) = \prod_\alpha \left(1 - x^\alpha\right)^{-1}$$
$$B_\lambda(x) = \prod_\alpha \left(1 + x^\alpha\right),$$

where α ranges over all *distinct* permutations of $(\lambda_1, \lambda_2, \ldots)$. Find a formula for $\omega A_\lambda(x)$ and $\omega B_\lambda(x)$ in terms of $A_\mu(x)$'s and $B_\mu(x)$'s. For instance,

$$\begin{aligned} \omega A_1(x) &= B_1(x) \\ \omega A_{11}(x) &= A_2(x) A_{11}(x) \\ \omega A_2(x) &= A_2(x)^{-1}. \end{aligned}$$

In general, express the answer in terms of the coefficients $a_{\lambda\mu}$ defined in Exercise 7.9.

7.11. [2+] Let q be an indeterminate. Find the Schur function expansion of $\sum_{\mu \vdash n} q^{\ell(\mu)-1}\, m_\mu$.

7.12. [3−] Prove the converse to Proposition 7.10.5, i.e., if $\mu, \lambda \vdash n$ and $\mu \leq \lambda$ (dominance order), then $K_{\lambda\mu} \neq 0$.

7.13. **a.** [3−] Let $\lambda \vdash n$ and $\mu \vdash n$, with $\lambda \neq (n)$. Suppose that $\lambda_1 \neq \mu_1$ and $\lambda_1' \neq \mu_1'$. Show that $K_{\lambda\mu} = 1$ if and only if $\lambda = \langle (m+1)^m \rangle$ for some m (so $n = m(m+1)$), and $\mu = \langle m^{m+1} \rangle$. (Note that μ is assumed to be a *partition*, not just a composition.) Note that this result gives a complete characterization of when $K_{\lambda\mu} = 1$, since if $\lambda_1 = \mu_1$ then $K_{\lambda\mu} = K_{(\lambda_2, \lambda_3, \ldots), (\mu_2, \mu_3, \ldots)}$, while if $\lambda_1' = \mu_1' = \ell$ then $K_{\lambda\mu} = K_{(\lambda_1 - 1, \ldots, \lambda_\ell - 1), (\mu_1 - 1, \ldots, \mu_\ell - 1)}$.

b. [5−] Find a "reasonable" characterization of partitions λ, μ, ν for which the Littlewood–Richardson coefficient $c^\lambda_{\mu\nu}$ is equal to one.

7.14. **a.** [2] Show that the number of SSYTs of type (r, r, r) (i.e., with r 1's, r 2's, and r 3's, and no other parts) is equal to

$$\tfrac{1}{16}[4r^3 + 18r^2 + 28r + 15 + (-1)^r].$$

b. [3+] Fix $n \geq 1$. Show that there are polynomials $P_n(r)$ of degree $\binom{n}{2}$ and $Q_n(r)$ of degree $\binom{2\lfloor (n-1)/2 \rfloor}{2} - 1$ such that the number of SSYTs of type $(r, r, \ldots, r)$ (k times) is given by $P_n(r) + (-1)^r Q_n(r)$. (The most difficult part is the degree of $Q_n(r)$.)

7.15. **a.** [2+] Let p be prime. Let $M_p(n)$ denote the number of partitions λ of n such that the number f^λ of SYTs of shape λ is relatively prime to p. Let

$$n = \alpha_0 + \alpha_1 p + \alpha_2 p^2 + \cdots, \qquad 0 \leq \alpha_i \leq p - 1,$$

the base p expansion of n. Let $P(x) = \prod_{n \geq 1} (1 - x^n)^{-1}$. Show that

$$M_p(n) = \prod_{j \geq 0} \left(\text{coefficient of } x^{\alpha_j} \text{ in } P(x)^{p^j}\right).$$

b. [1+] Deduce from (a) that if $n = p^{k_1} + p^{k_2} + \cdots$ with $k_1 < k_2 < \cdots$, then $M_p(n) = p^{k_1 + k_2 + \cdots}$.

7.16. **a.** [3] Let $B_k = \sum_\lambda s_\lambda$, summed over all partitions with at most k parts. Let

$$c_i = \sum_{n=0}^{\infty} h_n h_{n+i}.$$

Show that

$$B_{2m} = \det(c_{i-j} + c_{i+j-1})_{1 \leq i,j \leq m}$$

$$B_{2m+1} = h \cdot \det(c_{i-j} + c_{i+j})_{1 \leq i,j \leq m},$$

where $h = \sum_{n \geq 0} h_n$.

b. [3−] Let $y_k(n)$ be the number of SYTs with n entries and at most k rows, and let C_n denote the Catalan number $\frac{1}{n+1}\binom{2n}{n}$. Deduce from (a) that

$$y_2(n) = \binom{n}{\lfloor n/2 \rfloor}$$

$$y_3(n) = \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{2i} C_i$$

$$y_4(n) = C_{\lfloor (n+1)/2 \rfloor} C_{\lceil (n+1)/2 \rceil}$$

$$y_5(n) = 6 \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{2i} C_i \frac{(2i+2)!}{(i+2)!(i+3)!}.$$

c. [2]–[3+] Give combinatorial proofs of the above formulas for $y_2(n), \ldots, y_5(n)$. Also give a simple symmetric function proof of the formula for $y_2(n)$, by considering the product $s_{\lfloor n/2 \rfloor} s_{\lceil n/2 \rceil}$.

d. [3] Let $R_k(x, y) = \sum_\lambda s_\lambda(x)s_\lambda(y)$, summed over all partitions with at most k parts. Let

$$A_i = \sum_{l=0}^{\infty} h_{l+i}(x)h_l(y).$$

Show that $R_k(x, y) = \det(A_{j-i})_{i,j=1}^k$.

e. [2+] Let $u_k(n)$ be the number of pairs of SYTs of the same shape with n entries each and at most k rows. Deduce from (c) that $u_2(n) = C_n$ and

$$u_3(n) = \frac{1}{(n+1)^2(n+2)} \sum_{k=0}^{n} \binom{2k}{k}\binom{n+1}{k+1}\binom{n+2}{k+1}. \qquad (7.166)$$

f. [2], [5−] Give combinatorial proofs of the above formulas for $u_2(n)$ and $u_3(n)$.

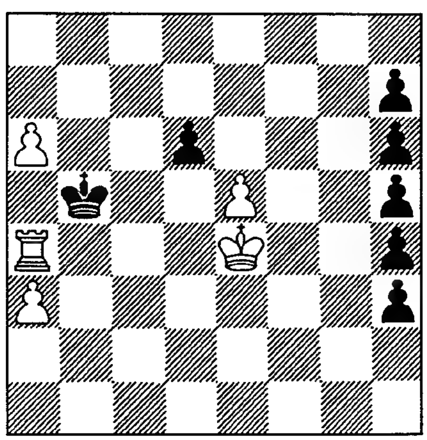
7.17. **a.** [3−] Let $W_i(n)$ be the number of ways to draw i diagonals in a convex n-gon such that no two diagonals intersect in their interiors. Give a bijective proof that $W_i(n)$ is the number of standard Young tableaux of shape $\langle (i+1)^2, 1^{n-i-3}\rangle$ (i.e., two parts equal to $i+1$ and $n-i-3$ parts equal to 1; when $i = 0$ there are $n-1$ parts equal to 1).

b. [2−] Deduce from (a) and the hook-length formula (Corollary 7.21.6) that

$$W_i(n) = \frac{1}{n+i}\binom{n+i}{i+1}\binom{n-3}{i},$$

in agreement with equation (6.74).

7.18. [3−] Solve the following chess problem:



Serieshelpmate in 25. How many solutions?

The definition of serieshelpmate is given in Exercise 6.23.

7.19. [3−] Let σ be a skew shape, regarded as a subset of $\mathbb{P}\times\mathbb{P}$. For instance, $32/1 = \{(1, 2), (1, 3), (2, 1), (2, 2)\}$. Let $p_\sigma(n)$ be the number of pairs (λ, μ) of partitions for which $\mu \subseteq \lambda$, $\mu \vdash n$, and λ/μ is a *translate* of σ. For instance, if

$\sigma = 32/1$ then $p_\sigma(6) = 3$, corresponding to $\lambda/\mu = 532/51,\ 43111/21111,\ 541/321$. Suppose that the smallest rectangle containing σ has r rows and s columns, and that t is the smallest integer for which $\sigma = \rho/\nu$ with $\nu \vdash t$. Show that

$$\sum_{n\geq 0} p_\sigma(n+t)q^n = \frac{[s-1]!\,[r-1]!}{[\infty]!\,[r+s-1]!},$$

where $[k]! = \prod_{i=1}^{k}(1-q^i)$.

7.20. **a.** [2] Let $\lambda \vdash n$. Show that $[\prod_i (m_i(\lambda)!)^{-1}]\langle p_1^n, h_\lambda\rangle$ is equal to the number of partitions of the set $[n]$ of type λ (i.e., with block sizes $\lambda_1, \lambda_2, \ldots$).

b. [2+] Let T be an SYT of shape $\lambda \vdash n$. For each entry of T not in the first column, let $f(i)$ be the number of entries j in the column immediately to the left of i and in a row not above i, for which $j < i$. Define $f(T) = \prod_i f(i)$, where i ranges over all entries of T not in the first column. For instance, if

$$T = \begin{matrix} 1\,3\,6\,8 \\ 2\,4\,7 \\ 5 \end{matrix}\ ,$$

then $f(3) = 2$, $f(4) = 1$, $f(6) = 2$, $f(7) = 1$, $f(8) = 2$, and $f(T) = 8$. Show that

$$\sum_T f(T) = \left[\prod_i (m_i(\lambda)!)^{-1}\right]\langle p_1^n, h_\lambda\rangle,$$

where T ranges over all SYTs of shape λ.

c. [3−] Generalize (b) to SSYTs T of a given shape λ and type μ (so (b) is the special case $\mu = \langle 1^n\rangle$).

7.21. [3] Let $\lambda \vdash n$. An assignment $u \mapsto a_u$ of the distinct integers $1, 2, \ldots, n$ to the squares $u \in \lambda$ is a *balanced tableau* of shape λ if for each $u \in \lambda$ the number a_u is the k-th largest number in the hook of u, where k is the leg length (number of squares directly below u, counting u itself) of the hook of u. For instance, the balanced tableaux of shape (3, 2) are

$$\begin{matrix} 421 & 423 & 425 & 435 & 321 \\ 53 & 51 & 31 & 21 & 54 \end{matrix}\ .$$

Let b^λ be the number of balanced tableaux of shape λ. Show that $b^\lambda = f^\lambda$, the number of SYTs of shape λ.

7.22. Let s_i denote the adjacent transposition $(i, i+1) \in \mathfrak{S}_n$, for $1 \leq i \leq n-1$. Let $w \in \mathfrak{S}_n$. It is easy to see that the smallest integer p for which w is a product of p adjacent transpositions is equal to $\ell(w)$, the number of inversions of w (defined in Section 1.3 and there denoted $i(w)$). A *reduced decomposition* of w is a sequence $a = (a_1, \ldots, a_p)$, where $p = \ell(w)$, such that $w = s_{a_1}\cdots s_{a_p}$. As usual, define the *descent set* $D(a) = \{i : 1 \leq i \leq p-1 \text{ and } a_i > a_{i+1}\}$; and write $\mathrm{co}(a)$ for the composition $\mathrm{co}(D(a)) \in \mathrm{Comp}(p)$, as defined in Section 7.19.

a. [1+] Let $R(w)$ denote the set of reduced decompositions of w, and $r(w) = \#R(w)$. Define the quasisymmetric function

$$F_w = \sum_{a \in R(w)} L_{\mathrm{co}(a)}.$$

Show that $r(w) = [x_1 x_2 \cdots x_p] F_w$. ($L_\alpha$ is defined in equation (7.89).)

b. [3] Show that $F_w \in \Lambda^p$, i.e., F_w is a homogeneous *symmetric* function of degree p. Hence if $F_w = \sum_{\lambda \vdash p} c_{w\lambda} s_\lambda$, then

$$r(w) = \sum_{\lambda \vdash p} c_{w\lambda} f^\lambda. \tag{7.167}$$

c. [3−] Define

$$r_i(w) = \#\{j \, : \, j < i \text{ and } a_j > a_i\}, \quad 1 \le i \le n$$
$$s_i(w) = \#\{j \, : \, j > i \text{ and } a_j < a_i\}, \quad 1 \le i \le n.$$

Thus $\sum_i r_i(w) = \sum_i s_i(w) = \ell(w)$. Let $\lambda(w)$ denote the partition obtained by arranging the numbers $r_1(w), \ldots, r_n(w)$ in descending order (and ignoring any 0's). Let $\mu(w)$ denote the *conjugate* to the partition $\mu'(w)$ obtained by arranging the numbers $s_1(w), \ldots, s_n(w)$ in descending order. Show that if $c_{w\nu} \neq 0$ then $\mu(w) \le \nu \le \lambda(w)$ (dominance order). Moreover, $c_{w,\mu(w)} = c_{w,\lambda(w)} = 1$. Hence F_w is a single Schur function s_ν (in which case $r(w) = f^\nu$) if and only if $\mu(w) = \lambda(w) = \nu$.

d. [3−] Show that $\lambda(w) = \mu(w)$ if and only if $w = w_1 \cdots w_n$ is 2143-*avoiding*, i.e., there do not exist $a < b < c < d$ such that $w_b < w_a < w_d < w_c$. 2143-avoiding permutations are also called *vexillary*, after the Latin word *vexillum* for "flag," because the Schubert polynomial indexed by a vexillary permutation is a flag Schur function; we will not define "Schubert polynomial" and "flag Schur function" here.

e. [3] Let $v(n)$ be the number of vexillary permutations in $\mathfrak{S}_n$. Show that

$$v(n) = \sum_{\substack{\lambda \vdash n \\ \ell(\lambda) \le 3}} (f^\lambda)^2.$$

A more explicit formula for $v(n)$ then follows from equation (7.166).

f. [2−] Let $w_0 = n, n-1, \ldots, 1$, the permutation in $\mathfrak{S}_n$ with the maximum number of inversions. Deduce from (c) that the number of reduced decompositions of w_0 is given by

$$r(w_0) = \frac{\binom{n}{2}!}{1^{n-1}\, 3^{n-2} \cdots (2n-3)^1} = \binom{n}{2}! \prod_{1 \le i < j \le n} \frac{2}{i+j-1}.$$

g. [3+] Let $p = \binom{n}{2}$. Show that

$$\sum_{(a_1, \ldots, a_p) \in R(w_0)} (x + a_1) \cdots (x + a_p) = \binom{n}{2}! \prod_{1 \le i < j \le n} \frac{2x + i + j - 1}{i + j - 1}. \tag{7.168}$$

Note that taking the coefficient of x^p on both sides gives (f). Moreover, setting $x = 0$ yields

$$\sum_{(a_1,\ldots,a_p)\in R(w_0)} a_1 \cdots a_p = \binom{n}{2}!. \qquad (7.169)$$

h. [3] Show that $c_{w\lambda} \geq 0$ for all $w \in \mathfrak{S}_n$ and $\lambda \vdash \ell(w)$.

7.23. [3−] Let P be a finite graded poset of rank n. A *symmetric chain decomposition* of P is a partition of P into chains $x_i < x_{i+1} < \cdots < x_{n-i}$ such that x_j has rank j. Let M denote the (finite) multiset $\{1^{a_1}, 2^{a_2}, \ldots, k^{a_k}\}$, and let B_M denote the set of all submultisets of M, ordered by inclusion. (Thus B_M is just a product of chains of lengths $a_1, a_2, \ldots, a_k$.) Associate with each submultiset $N = \{1^{b_1}, \ldots, 1^{b_k}\} \in B_M$ the two-line array w_N whose first line consists of a_1 1's, then a_2 2's, etc., and whose second line has $a_i - b_i$ 0's followed by b_i 1's below the i's in the first line. Call two submultisets N and N' *equivalent* if, when the RSK algorithm is applied to w_N and $w_{N'}$, the same second tableau is obtained. This definition of equivalence partitions B_M into equivalence classes. Show that they form a symmetric chain decomposition of B_M.

7.24. **a.** [1] Let $U : \Lambda \to \Lambda$ and $D : \Lambda \to \Lambda$ be linear transformations defined by $U(f) = p_1 f$ and

$$D(f) = \frac{\partial}{\partial p_1} f,$$

where $\partial/\partial p_1$ is applied to f written as a polynomial in the p_i's. Show that $DU - UD = I$, the identity operator.

b. [1] Show that $DU^k = kU^{k-1} + U^k D$.

c. [2] Deduce from (a) and (b) that if $\ell \in \mathbb{N}$ then

$$(U + D)^\ell = \sum_{\substack{i+j\leq \ell \\ r:=(\ell-i-j)/2\in\mathbb{N}}} \frac{\ell!}{2^r r!\, i!\, j!} U^i D^j.$$

d. [2+] An *oscillating tableau* (or *up–down tableau*) of shape λ and length ℓ is a sequence $\emptyset = \lambda^0, \lambda^1, \ldots, \lambda^\ell = \lambda$ of partitions such that for all $1 \leq i \leq \ell - 1$, the diagram of λ^i is obtained from that of λ^{i-1} by either adding one square or removing one square. (If we add a square each time, then $\ell = |\lambda|$ and we have an SYT of shape λ.) Clearly if such an oscillating tableau exists, then $\ell = |\lambda| + 2r$ for some $r \in \mathbb{N}$. Deduce from (c) that the number $\tilde{f}^\lambda_\ell$ of oscillating tableaux of shape λ and length $\ell = |\lambda| + 2r$ is given by

$$\tilde{f}^\lambda_\ell = \frac{\ell!\, f^\lambda}{(\ell - 2r)!\, r!\, 2^r}.$$

e. [3−] Give a bijective proof of (d).

7.25. **a.** [3] Let $f_{2k}(n)$ be the number of ways to choose the diagram of a partition λ of n, then add or remove one square at a time for a total of $2k$ times, always

keeping the diagram of a partition, and ending back at λ. Show that

$$\sum_{n\geq 0} f_{2k}(n)q^n = \frac{(2k)!}{2^k k!}\left(\frac{1+q}{1-q}\right)^k \prod_{i\geq 1}(1-q^i)^{-1}.$$

(Note that the case $n = 0$, obtained by setting $q = 0$, coincides with the case $\lambda = \emptyset$ of Exercise 7.24(d).)

b. [3−] Let $g_{2k}(n)$ be the number of ways to choose the diagram of a partition λ of n, then remove a square, then add a square, then remove a square, etc., for a total of k additions and k removals, always keeping the diagram of a partition, and ending back at λ. For instance, $g_0(n) = p(n)$, the number of partitions of n. Show that

$$g_{2k}(n) = \sum_{j=0}^{n}(p(j) - p(j-1))(n-j)^k.$$

7.26. [3+] Given $u = (i, j) \in \lambda \vdash n$, define the *arm length* $a(u) = \lambda_i - j$ and the *leg length* $\ell(u) = \lambda'_j - i$. Note that the hook length $h(u)$ satisfies $h(u) = a(u) + \ell(u) + 1$. Prove the identity

$$\sum_{\lambda\vdash n} \frac{\left(\sum_{(i,j)\in\lambda} t^{i-1}q^{j-1}\right)\prod_{(i,j)\in\lambda\setminus(1,1)}(1-t^{i-1}q^{j-1})}{\prod_{u\in\lambda}\left(1-t^{-\ell(u)}q^{1+a(u)}\right)\left(1-t^{1+\ell(u)}q^{-a(u)}\right)} = \frac{1}{(1-t)(1-q)}.$$

7.27. [3] Prove the following identities (combinatorially if possible). Here α and β are fixed partitions.

(a) $$\sum_{n\geq 0}\sum_{\substack{\mu\subseteq\lambda\\ |\lambda/\mu|=n}} f^{\lambda/\mu}\frac{q^{|\mu|}t^n}{n!} = \exp\left(\frac{t}{1-q}+\frac{t^2}{2(1-q^2)}\right)\cdot\prod_{i\geq 1}(1-q^i)^{-1}$$

(b) $$\sum_{n\geq 0}\sum_{\substack{\mu\subseteq\lambda\\ |\lambda/\mu|=n}} (f^{\lambda/\mu})^2\frac{q^{|\mu|}t^n}{n!} = \frac{1}{1-\frac{t}{1-q}}\prod_{i\geq 1}(1-q^i)^{-1}$$

(c) $$\sum_{\lambda} s_{\lambda/\alpha}(x)s_{\lambda/\beta}(y) = \left(\prod_{i,j}(1-x_iy_j)^{-1}\right)\sum_{\mu} s_{\beta/\mu}(x)s_{\alpha/\mu}(y)$$

(d) $$\sum_{\lambda} s_{\lambda/\alpha'}(x)s_{\lambda'/\beta}(y) = \left(\prod_{i,j}(1+x_iy_j)\right)\sum_{\mu} s_{\beta'/\mu}(x)s_{\alpha/\mu'}(y)$$

(e) $$\sum_{\lambda} s_{\lambda/\alpha}(x) = \left(\prod_{i}(1-x_i)^{-1}\prod_{i<j}(1-x_ix_j)^{-1}\right)\sum_{\mu} s_{\alpha/\mu}(x)$$

(f) $$\sum_{\mu\subseteq\lambda} s_{\lambda/\mu}(x)s_{\lambda/\mu}(y)q^{|\mu|} = \prod_{n\geq 1}\left((1-q^n)\prod_{i,j}(1-x_iy_jq^{n-1})\right)^{-1}$$

(g) $$\sum_{\mu \subseteq \lambda} s_{\lambda/\mu}(x) s_{\lambda'/\mu'}(y) q^{|\mu|} = \prod_{n \geq 1} (1-q^n)^{-1} \prod_{i,j} (1 + x_i y_j q^{n-1})$$

(h) $$\sum_{\mu \subseteq \lambda} s_{\lambda/\mu}(x) q^{|\mu|} = \prod_{n \geq 1} \left((1-q^n) \prod_{i<j} (1 - x_i x_j q^{n-1}) \prod_i (1 - x_i q^{n-1}) \right)^{-1}$$

7.28. **a.** [3−] Suppose that in the RSK algorithm $A \xrightarrow{\text{RSK}} (P, Q)$, the matrix A is symmetric (so $P = Q$). Show that $\operatorname{tr}(A)$ is the number of columns of P of odd length.

b. [2−] Verify the identity

$$\prod_i (1 - qx_i)^{-1} \cdot \prod_{i<j} (1 - x_i x_j)^{-1} = \sum_\lambda q^{c(\lambda)} s_\lambda(x), \tag{7.170}$$

where $c(\lambda)$ denotes the number of parts of λ' that are odd.

c. [1] Deduce that

$$\prod_{i<j} (1 - x_i x_j)^{-1} = \sum_\mu s_{(2\mu)'}(x),$$

where $2\mu = (2\mu_1, 2\mu_2, \ldots)$.

d. [2−] Fix $k \geq 0$. Evaluate the sum $a(n,k) = \sum_\lambda f^\lambda$, where λ ranges over all partitions of n with k odd parts.

e. [2] What identity results when we apply ω to (7.170)?

7.29. **a.** [3−] Show that

$$\prod_i (1 - x_i) \cdot \prod_{i<j} (1 - x_i x_j) = \sum_\lambda (-1)^{\frac{1}{2}(|\lambda| + \operatorname{rank}(\lambda))} s_\lambda, \tag{7.171}$$

where λ ranges over all self-conjugate partitions.

b. [3−] Show that

$$\prod_i \left(1 + x_i^2\right) \cdot \prod_{i<j} (1 + x_i x_j) = \sum_\lambda s_\lambda,$$

where λ ranges over all partitions whose Frobenius notation (as defined in Exercise 7.39) has the form

$$\lambda = (\alpha_1 + 1 \ \cdots \alpha_r + 1 \mid \alpha_1 \cdots \alpha_r).$$

c. [3−] Show that

$$\prod_i (1 + x_i)^{-1} \cdot \prod_{i \leq j} (1 + x_i x_j)^{-1} = \sum_\lambda (-1)^{\frac{|\lambda| + o(\lambda)}{2}} s_\lambda,$$

where $o(\lambda)$ is the number of odd parts of λ, and λ runs over all partitions satisfying

$$\begin{aligned} \lambda_i - \lambda_{i+1} &\equiv 0, 1 \pmod 4, & \lambda_i \text{ even} \\ \lambda_i - \lambda_{i+1} &\equiv 1, 2 \pmod 4, & \lambda_i \text{ odd}. \end{aligned}$$

7.30. **a.** [2] Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ and $\mu = (\mu_1, \ldots, \mu_n)$ be partitions of length at most n related by

$$\lambda_i + n - i = d(\mu_i + n - i), \qquad 1 \le i \le n,$$

for some fixed $d \in \mathbb{P}$. Show that

$$s_\lambda(x_1, \ldots, x_n) = s_\mu\left(x_1^d, \ldots, x_n^d\right) \prod_{1 \le i < j \le n} \frac{x_i^d - x_j^d}{x_i - x_j}.$$

b. [1+] Suppose that $\lambda = (d(n-1), d(n-2), \ldots, d, 0)$. Deduce from (a) that

$$s_\lambda(x_1, \ldots, x_n) = \prod_{1 \le i < j \le n} \left(x_i^{d-1} + x_i^{d-2} x_j + x_i^{d-3} x_j^2 + \cdots + x_j^{d-1}\right).$$

c. [3−] It follows from (b) that the number of SSYTs of shape $\lambda = (d(n-1), d(n-2), \ldots, d)$ and type $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ (where $\alpha_1 + \cdots + \alpha_n = d\binom{n}{2}$) is equal to the number of ways of orienting the edges of the graph on the vertex set $\{1, 2, \ldots, n\}$ with $d - 1$ edges between any two distinct vertices, such that vertex i has outdegree α_i for $1 \le i \le n$. Give a direct combinatorial proof.

7.31. [3] Let p be a prime, and let A_p denote the matrix $[\zeta^{jk}]_{j,k=0}^{p-1}$, where $\zeta = e^{2\pi i/p}$. Show that every minor of A_p is nonzero. Equivalently, every square submatrix B of A_p is invertible. (Hint. Use Theorem 7.15.1.)

7.32. **a.** [2+] Let λ and μ be partitions of length at most n. Show that

$$s_\lambda(q^{\mu_1+n-1}, q^{\mu_2+n-2}, \ldots, q^{\mu_n}) = s_\mu(q^{\lambda_1+n-1}, q^{\lambda_2+n-2}, \ldots, q^{\lambda_n}) \times \prod_{1 \le i < j \le n} \frac{1 - q^{\lambda_i - \lambda_j + j - i}}{1 - q^{\mu_i - \mu_j + j - i}}. \qquad (7.172)$$

b. [2] Deduce from (a) that

$$s_\lambda(1, q, q^2, \ldots, q^{n-3}, q^{n-2}, q^n) = \frac{\sum_i q^{\lambda_i + n - i}}{1 + q + \cdots + q^{n-1}} s_\lambda(1, q, \ldots, q^{n-1})$$

$$s_\lambda(1, q^2, q^3, \ldots, q^{n-2}, q^{n-1}, q^{n+1}) = \frac{q^{\binom{n+1}{2}}\left(n - 1 + \sum_{i \ne j} q^{\lambda_i - \lambda_j + j - i}\right)}{(1 + q + \cdots + q^{n-2})(1 + q + \cdots + q^n)} s_\lambda(1, q, \ldots, q^{n-1}).$$

7.33. **a.** [2+] Let $\delta = (n-1, n-2, \ldots, 1)$. Let $t(n)$ denote the number of distinct monomials appearing in $s_\delta(x_1, \ldots, x_n)$, i.e., the number of sequences $\alpha = (\alpha_1, \ldots, \alpha_n)$ for which $K_{\delta\alpha} \ne 0$. For example, $t(3) = 7$. Show that $t(n)$ is equal to the number of labeled forests on n vertices, which by equation (5.42) and Proposition 5.3.2 is given by

$$\sum_{n \ge 0} t(n) \frac{x^n}{n!} = \exp \sum_{j \ge 1} j^{j-2} \frac{x^j}{j!}.$$

b. [3−] Let $k \in \mathbb{P}$. Generalize (a) to $s_{k\delta}(x_1, \dots, x_n)$.

c. [5−] Can anything be said in general about the number of distinct monomials in $s_\lambda(x_1, \dots, x_n)$ for arbitrary λ?

7.34. [3−] Let λ and μ be partitions of length at most n. Show that in the ring Λ_n (i.e., using only n variables), we have

$$s_\lambda s_\mu = \det\left(h_{\lambda_i+\mu_{n+1-j}-i+j}\right)_{i,j=1}^n.$$

7.35. a. [2] If R is a ring, then an additive group homomorphism $D : R \to R$ is called a *derivation* if $D(fg) = (Df)g + f(Dg)$ for all $f, g \in R$. Show that the linear transformation $\Lambda \to \Lambda$ defined by $D(s_\lambda) = s_{\lambda/1}$ is a derivation.

b. [2] Show that the bilinear operation $[f, g]$ on Λ given by $[s_\lambda, s_\mu] = s_{\lambda/1}s_\mu - s_\lambda s_{\mu/1}$ defines a Lie algebra structure on Λ. (In other words, verify the Jacobi identity $[f, [g, h]] + [g, [h, f]] + [h, [f, g]] = 0$.)

c. [3−] Let $\rho_m = (m, m-1, \dots, 1)$. Show that

$$\left[s_{\rho_{n+1}}, s_{\rho_{n-1}}\right] = s_{\rho_n}^2.$$

7.36. [2] Let $D_\mu : \Lambda \to \Lambda$ be the linear transformation given by $D_\mu(s_\lambda) = s_{\lambda/\mu}$. Show that $D_\mu D_\nu = D_\nu D_\mu$.

7.37. a. [2+] Let $a_\delta = \prod_{1\le i<j\le n}(x_i - x_j)$, as in equation (7.53). Write down a formula that expresses a_δ^2 in terms of the power sums $p_i(x_1, \dots, x_n)$, $1 \le i \le 2n-2$. (You don't need to compute explicitly the coefficients in the expansion of a_δ^2 in terms of power sums; just some formula involving only power sums is wanted.)

b. [3−] Let $a_\delta^2 = \sum_{\lambda\vdash n(n-1)} c_\lambda s_\lambda(x_1, \dots, x_n)$, where $c_\lambda \in \mathbb{Z}$. Show that if λ is the partition $\langle (n-1)^n\rangle$, then

$$c_\lambda = (-1)^{\binom{n}{2}} \cdot 1 \cdot 3 \cdot 5 \cdots (2n-1).$$

c. [3] More generally, if $\lambda = \langle (n+i-1)^{n-i}, (i-1)^i\rangle$, $1 \le i \le n$, then

$$c_\lambda = (-1)^{\frac{1}{2}(n-1)(n-2i)}[1 \cdot 3 \cdot 5 \cdots (2i-1)] \cdot [1 \cdot 3 \cdot 5 \cdots (2(n-i)-1)].$$

d. [3] Suppose that $\lambda = \mu + \langle (n-2)^n\rangle = (\mu_1 + n - 2, \dots, \mu_n + n - 2)$, so $\mu \vdash n$. Show that

$$c_\lambda = (-1)^{\binom{n}{2}} f^\lambda \prod_{s\in\lambda}(1 - 2c(s)),$$

where $c(s)$ is the content of s, and f^λ is the number of SYTs of shape λ.

e. [3−] Show that if $\lambda \ne 2\delta = 2(n-1, n-2, \dots, 1)$, then $c_\lambda \equiv 0 \pmod 3$.

7.38. a. [3] Fix $0 \le k \le \binom{n}{2}$, and let $\ell(w)$ denote the number of inversions of the permutation $w \in \mathfrak{S}_n$. Let λ and μ be partitions of length at most n, with $\mu \subseteq \lambda$. Define the symmetric function

$$t_{\lambda/\mu,k} = (-1)^k \sum_{\substack{w\in\mathfrak{S}_n\\ \ell(w)\ge k}} \varepsilon_w h_{\lambda+\delta-w(\mu+\delta)}.$$

Thus $t_{\lambda/\mu,k}$ is a "truncation" of the Jacobi–Trudi expansion (7.69) of $s_{\lambda/\mu}$. Show that $t_{\lambda/\mu,k}$ is s-positive.

b. [5−] Is there a "nice" combinatorial interpretation of the scalar product $\langle t_{\lambda/\mu,k}, s_\nu\rangle$?

7.39. [3−] Let λ be a partition of rank r. For $1 \le i \le r$ define $\alpha_i = \lambda_i - i$ and $\beta_i = \lambda'_i - i$. The *Frobenius notation* for λ is the array

$$\lambda = (\alpha_1\, \alpha_2\, \cdots\, \alpha_r \mid \beta_1\, \beta_2\, \cdots\, \beta_r). \tag{7.173}$$

For instance,

$$(7,7,5,4,4,2,1,1) = (6\,5\,2\,0 \mid 7\,4\,2\,1).$$

Note that if $a, b \in \mathbb{N}$, then $(a \mid b)$ is the Frobenius notation for the hook shape $\langle a+1, 1^b\rangle$. It is easy to see that any array (7.173) of integers satisfying $\alpha_1 > \alpha_2 > \cdots > \alpha_r \ge 0$ and $\beta_1 > \beta_2 > \cdots > \beta_r \ge 0$ is the Frobenius notation of a unique partition of rank r. Show that if $\lambda = (\alpha_1 \cdots \alpha_r \mid \beta_1 \cdots \beta_r)$, then

$$s_\lambda = \det\left(s_{(\alpha_i \mid \beta_j)}\right)_{i,j=1}^r.$$

7.40. [3−] Let u be a square of (the diagram of) the partition λ. Given $(i,j) \in \lambda$, let $B(i,j)$ be the border strip of λ whose top square is in row i of λ and whose bottom square is in column j of λ. Let r be the rank of λ. Show that

$$s_\lambda = \det\left(s_{B(i,j)}\right)_{i,j=1}^r.$$

7.41. [2+] Let $\ell(\lambda) \le m$ and $\lambda_1 \le n$. Define

$$\tilde{\lambda} = (n - \lambda_m, n - \lambda_{m-1}, \ldots, n - \lambda_1).$$

Give simple algebraic and combinatorial proofs that

$$(x_1 x_2 \cdots x_m)^n s_\lambda\left(x_1^{-1}, \ldots, x_m^{-1}\right) = s_{\tilde{\lambda}}(x_1, \ldots, x_m).$$

7.42. [2+] Show that

$$\prod_{i=1}^m \prod_{j=1}^n (x_i + y_j) = \sum_\lambda s_\lambda(x) s_{\tilde{\lambda}'}(y),$$

summed over all partitions λ with $\ell(\lambda) \le m$ and $\lambda_1 \le n$, where $\tilde{\lambda}$ is defined in Exercise 7.41 above.

7.43. [3−] Let $\psi : \Lambda \to \mathbb{Q}[t]$ be the specialization (homomorphism) defined by $\psi(p_n) = 1 - (-t)^n$, $n > 0$. Show that

$$\psi(s_\lambda) = \begin{cases} t^k(1+t), & \lambda = \langle n-k, 1^k\rangle,\ 0 \le k \le n-1 \\ 0 & \text{otherwise.} \end{cases}$$

7.44. Define a specialization $\xi : \Lambda \to \mathbb{Q}[t]$ by

$$\xi(e_n) = \frac{(1-t)^{n-1}}{n!}, \qquad n > 0.$$

Let $\lambda \vdash n$ in (b)–(d) below.

a. [2+] Show that

$$\xi(h_n) = \frac{A_n(t)}{t \cdot n!},$$

where $A_n(t)$ denotes an Eulerian polynomial (Section 1.3).

b. [3−] Show that

$$\xi(p_\lambda) = \frac{z_\lambda}{n!} \sum_{\substack{w \in \mathfrak{S}_n \\ \rho(w)=\lambda}} t^{e(w)},$$

where $e(w) = \#\{i : w(i) > i\}$, the number of excedances of w.

c. [2+] Show that

$$\xi(s_\lambda) = \frac{1}{n!} \sum_{\substack{w \in \mathfrak{S}_n \\ c(w) \geq \mathrm{rank}(\lambda)}} \chi^\lambda(w) t^{e(w)},$$

where $c(w)$ is the number of cycles of w.

d. [2+] Show that

$$\xi(s_\lambda) = \frac{1}{n!} \sum_{\mu \vdash n} \binom{n}{\mu_1, \mu_2, \dots} (K^{-1})_{\mu\lambda'} (1-t)^{n-\ell(\mu)},$$

where K^{-1} denotes the inverse of the Kostka matrix $(K_{\lambda\mu})$.

7.45. [3−] Suppose that $n = ab$, where $a, b \in \mathbb{P}$. If f is a symmetric function of degree n, then let $T_a(f)$ be the symmetric function obtained from f by expanding f in terms of monomials, dividing the exponents of these monomials by a, and then throwing away all terms whose exponents are not all integers. Thus $T_a(f)$ is a symmetric function of degree b. (For instance, $T_a(p_n) = p_b$, and $T_2(p_1^4) = m_2 + 6m_{11}$.) Show that if $\lambda \vdash n$, then $T_a(s_\lambda)$ is s-positive.

7.46. [3−] Recursively define symmetric functions q_n by

$$\sum_{\lambda \vdash n} q_\lambda = s_n,$$

where $q_\lambda = q_{\lambda_1} q_{\lambda_2} \cdots$. Show that for $n \geq 2$, the symmetric function $-q_n$ is s-positive.

7.47. Let G be a graph (without loops or multiple edges) on the d-element vertex set V. A *proper coloring* of G is a map $\kappa : V \to \mathbb{P}$ such that if $\{u, v\}$ is an edge of G, then $\kappa(u) \neq \kappa(v)$. Define $x^\kappa = \prod_i x_i^{\#\kappa^{-1}(i)}$, a monomial of degree d. Let $X_G = \sum_\kappa x^\kappa$, summed over all proper colorings of G. Thus the coefficient of $x_1^{a_1} x_2^{a_2} \cdots$ in X_G is the number of proper colorings of G such that a_i vertices are colored i for all $i \geq 1$. Clearly $X_G \in \Lambda^d$.

a. [1] Show that $X_G(1^n) = \chi_G(n)$, where χ_G denotes the chromatic polynomial of G (defined in Exercise 3.44). For this reason X_G is called the *chromatic symmetric function* of G.

b. [5] If T and T' are nonisomorphic trees, then is it true that $X_T \neq X_{T'}$?

c. [2−] A *stable partition* of G is a partition π of V such that every block B of π is *stable* (or *independent*), i.e., no two vertices of B are connected by an edge. Given a partition $\lambda = \langle 1^{r_1} 2^{r_2} \cdots \rangle$ of d, define the *augmented monomial symmetric function* $\tilde{m}_\lambda$ by $\tilde{m}_\lambda = r_1! \, r_2! \cdots m_\lambda$. Show that

$$X_G = \sum_{\lambda \vdash G} a_\lambda \tilde{m}_\lambda,$$

where a_λ is the number of stable partitions of G of type λ (i.e., with block sizes $\lambda_1, \lambda_2, \ldots$).

d. [2+] A *connected partition* of G is a partition π of V such that the restriction of G to every block of π is connected. Let L_G denote the lattice of all connected partitions of G ordered by refinement (as defined in Exercise 3.44). Show that

$$X_G = \sum_{\pi \in L_G} \mu(\hat{0}, \pi) p_{\mathrm{type}(\pi)}, \qquad (7.174)$$

where μ denotes the Möbius function of L_G.

e. [2−] Deduce from equation (7.174) and Proposition 3.10.1 that ωX_G is *p-positive.*

f. [3−] Show that X_G is *L-positive*, i.e., a nonnegative linear combination of the fundamental quasisymmetric functions L_α (defined in equation (7.89)), where $\alpha \in \mathrm{Comp}(d)$.

g. [3] let $X_G = \sum_{\lambda \vdash d} c_\lambda e_\lambda$, and fix $k \in \mathbb{P}$. Show that the integer $\sum_{\substack{\lambda \vdash d \\ \ell(\lambda) = k}} c_\lambda$ is equal to the number $\mathrm{sink}(G, k)$ of acyclic orientations of G with exactly k sinks. (A *sink* is a vertex u with no edge $u \to v$. In particular, an isolated vertex of G is a sink in any acyclic orientation of G.)

h. [3] Let P be a d-element poset, and let $\mathrm{inc}(P)$ denote its *incomparability graph*, i.e., the vertices of $\mathrm{inc}(P)$ are the elements of P, with u and v connected by an edge if u and v are incomparable in P. A *P-tableau* of shape $\lambda \vdash d$ is a map $\tau : P \to \mathbb{P}$ satisfying: (i) If $\tau(u) = \tau(v)$ then $u \le v$ or $v \le u$ (in other words, τ is a proper coloring of $\mathrm{inc}(P)$), (ii) $\#\tau^{-1}(i) = \lambda_i$ for all i, and (iii) if $\tau^{-1}(i) = \{u_1, u_2, \ldots, u_{\lambda_i}\}$ with $u_1 < u_2 < \cdots < u_{\lambda_i}$ and $\tau^{-1}(i+1) = \{v_1, v_2, \ldots, v_{\lambda_{i+1}}\}$ with $v_1 < v_2 < \cdots < v_{\lambda_{i+1}}$, then for all i and all $1 \le j \le \lambda_{i+1}$ we require that $v_j \not< u_j$. Let f_P^λ denote the number of P-tableaux of shape λ. (Note that if P is a chain, then $f_P^\lambda = f^\lambda$, the number of SYTs of shape λ.) Define P to be $(\mathbf{3}+\mathbf{1})$*-free* if it contains no induced subposet isomorphic to $\mathbf{3}+\mathbf{1}$ (the disjoint union of a three-element chain and a one-element chain). Show that if P is $(\mathbf{3}+\mathbf{1})$-free, then

$$X_{\mathrm{inc}(P)} = \sum_{\lambda \vdash d} f_P^\lambda s_\lambda.$$

i. [2+] Let P be a $(\mathbf{3}+\mathbf{1})$-free poset, and let c_i denote the number of i-element chains in P (so in particular $c_0 = 1$). Deduce from (h) and Exercise 7.91(e) that every zero of the polynomial $C(t) = \sum c_i t^i$ is real.

j. [5] Suppose that P is a $(\mathbf{3}+\mathbf{1})$-free poset. Is it true that X_G is e-positive?

k. [3−] Let P_d be a d-element path. Show that

$$\sum_{d \ge 0} X_{P_d} \cdot t^d = \frac{\sum_{i \ge 0} e_i t^i}{1 - \sum_{i \ge 1}(i-1) e_i t^i}. \qquad (7.175)$$

In particular, X_{P_d} is e-positive (a special case of (j)). Similarly, let C_d be a

d-vertex cycle. Show that

$$\sum_{d\geq 2} X_{C_d} \cdot t^d = \frac{\sum_{i\geq 0} i(i-1)e_i t^i}{1-\sum_{i\geq 1}(i-1)e_i t^i}.$$

l. [3−] Show that if the complement of G is triangle-free (equivalently, G contains no stable 3-element set of vertices), then X_G is e-positive.

m. [5] Suppose that G is *clawfree*, i.e., has no induced subgraph isomorphic to the complete bipartite graph $K_{1,3}$. Is it true that X_G is s-positive?

7.48. Let P be a finite graded poset of rank n with $\hat{0}$ and $\hat{1}$. Define a formal power series F_P in the variables $x_1, x_2, \ldots$ by the formula

$$F_P = \sum_{\hat{0}=t_0\leq t_1\leq\cdots\leq t_{k-1}<t_k=\hat{1}} x_1^{\rho(t_0,t_1)} x_2^{\rho(t_1,t_2)} \cdots x_k^{\rho(t_{k-1},t_k)}, \qquad (7.176)$$

where $\rho(s,t)$ denotes the rank (length) of the interval $[s,t]$. (The sum ranges over all multichains from $\hat{0}$ to $\hat{1}$ of all possible lengths $k\geq 1$ such that $\hat{1}$ occurs with multiplicity one.)

a. [2] Note that F_P is a homogeneous quasisymmetric function of degree n. Show that

$$F_P = \sum_{\gamma\in \mathrm{Comp}(n)} \beta_P(S_\gamma) L_\gamma,$$

where (i) $\beta_P(S_\gamma)$ is the rank-selected Möbius invariant (now called the *flag h-vector*) of P, as defined in Section 3.12, (ii) S_γ is the subset of $[n-1]$ associated with γ, as defined in Section 7.19, and (iii) L_γ is given by (7.89).

b. [2+] Define

$$\bar{F}_P = \sum_{\hat{0}=t_0\leq t_1\leq\cdots\leq t_{k-1}<t_k=\hat{1}} \mu(t_0,t_1)\mu(t_1,t_2)\cdots\mu(t_{k-1},t_k)$$
$$\times\, x_1^{\rho(t_0,t_1)} x_2^{\rho(t_1,t_2)} \cdots x_k^{\rho(t_{k-1},t_k)},$$

where μ denotes the Möbius function of P. Show that

$$\bar{F}_P = (-1)^n \sum_{\gamma\in \mathrm{Comp}(n)} \beta_P(\bar{S}_\gamma) L_\gamma,$$

where $\bar{S}_\gamma = [n-1] - S_\gamma$. Deduce that if $F_P \in \Lambda^n$, then $\bar{F}_P = (-1)^n \omega F_P$.

c. [2+] Define P to be *locally rank-symmetric* if every interval of P is rank-symmetric, i.e., has the same number of elements of rank i as of corank i for all i. For instance, if P is *locally self-dual* (i.e., every interval is self-dual), then P is locally rank-symmetric. Show that if P is locally rank-symmetric, then $F_P \in \Lambda^n$.

d. [2] Let $P = (\boldsymbol{\mu_1+1})\times\cdots\times(\boldsymbol{\mu_\ell+1})$, a product of chains of cardinalities $\mu_1+1,\ldots,\mu_\ell+1$. Show that P is locally self-dual, and that $F_P = h_\mu$.

e. [3+] Let P be the lattice of subgroups of a finite abelian p-group G of type μ. Show that P is locally-self dual, and that

$$F_P = \sum_{\lambda\vdash n} \tilde{K}_{\lambda\mu}(p) s_\lambda,$$

where $\tilde{K}_{\lambda\mu}(p)$ is a polynomial in p with nonnegative integer coefficients satisfying $\tilde{K}_{\lambda\mu}(1) = K_{\lambda\mu}$ (a Kostka number). (The most difficult part is the nonnegativity of the coefficients of $\tilde{K}_{\lambda\mu}(p)$.)

f. [3−] Let $P = \mathrm{NC}_{n+1}$, the lattice of noncrossing partitions of $[n+1]$, as defined in Exercises 3.68 and 5.35. Show that NC_{n+1} is locally self-dual, and that

$$\begin{aligned}
F_{\mathrm{NC}_{n+1}} &= \sum_{\lambda\vdash n}(n+1)^{\ell(\lambda)-1}\varepsilon_\lambda z_\lambda^{-1}p_\lambda \\
&= \sum_{\lambda\vdash n}\frac{1}{n+1}s_{\lambda'}(1^{n+1})s_\lambda \\
&= \sum_{\lambda\vdash n}\frac{1}{n+1}\left[\prod_i\binom{n+1}{\lambda_i}\right]m_\lambda \\
&= \sum_{\lambda\vdash n}\frac{n(n-1)\cdots(n-\ell(\lambda)+2)}{m_1(\lambda)!\cdots m_n(\lambda)!}e_\lambda \\
\omega F_{\mathrm{NC}_{n+1}} &= \sum_{\lambda\vdash n}\frac{1}{n+1}\left[\prod_i\binom{\lambda_i+n}{\lambda_i}\right]m_\lambda.
\end{aligned}$$

Here $m_i(\lambda)$ denotes the number of parts of λ equal to i. Show also that

$$\begin{aligned}
F_{\mathrm{NC}_{n+1}} &= \frac{1}{n+1}[t^n]E(t)^{n+1} \\
\sum_{n\geq 0}F_{\mathrm{NC}_{n+1}}t^{n+1} &= (tH(-t))^{\langle -1\rangle},
\end{aligned}$$

where $E(t) = \sum_{n\geq 0}e_nt^n$, $H(t) = \sum_{n\geq 0}h_nt^n$, and $^{\langle -1\rangle}$ denotes compositional inverse with respect to the variable t.

g. [3−] Let $m, n \in \mathbb{N}$. Define the *shuffle poset* (or *poset of shuffles*) W_{mn} as follows. Let $A = \{a_1, \ldots, a_m\}$ and $B = \{b_1, \ldots, b_n\}$ be two ordered alphabets. The elements of W_{mn} consist of all shuffles of subwords of the words $a_1\cdots a_m$ and $b_1\cdots b_n$, i.e, words whose restriction to the letters in A is a subword of $a_1\cdots a_m$, and similarly for B. Some examples of elements of W_{mn} are $\emptyset$ (the empty word), $b_2b_4a_3b_6a_4a_7b_7$, and $a_4a_5a_8b_1$. Define v to cover u in W_{mn} if v can be obtained from u either by deleting an element from A or inserting an element of B. Thus W_{mn} has minimum element $a_1a_2\cdots a_m$ and maximum $b_1b_2\cdots b_n$. Figure 7-16 shows the shuffle poset W_{21}, with $A = \{a, b\}$ and $B = \{x\}$. Show that W_{mn} is locally rank-symmetric (though not in general locally self-dual), and that

$$F_{W_{mn}} = \sum_{j\geq 0}\binom{m}{j}\binom{n}{j}e_1^{m+n-2j}e_2^j.$$

7.49. [3] Write F_n as short for the symmetric function $F_{\mathrm{NC}_{n+1}}$ of Exercise 7.48(f). Let $\psi : \Lambda \to \Lambda$ be the homomorphism defined by $\psi(h_n) = F_n$. Show that for

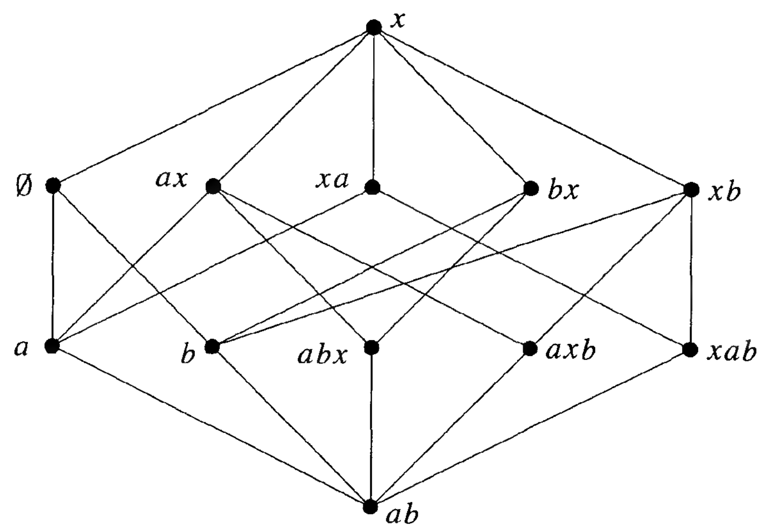


Figure 7-16. The shuffle poset W_{21}.

every skew shape λ/μ, the symmetric function $(-1)^{v(\lambda/\mu)}\psi(s_{\lambda/\mu})$ is s-positive, where $v(\lambda/\mu)$ is the number of nonzero entries below the main diagonal in the Jacobi–Trudi matrix for $s_{\lambda/\mu}$.

7.50. [2] Let $\lambda \vdash N$. Evaluate the sum

$$\frac{1}{N!}\sum_{w\in\mathfrak{S}_N}\chi^\lambda(w)n^{c(w)},$$

where $c(w)$ denotes the number of cycles of w. (Use the case $q = 1$ of Theorem 7.21.2.)

7.51. [2+] Show that if $\lambda \vdash N$ then

$$\binom{N}{2}\chi^\lambda(21^{N-2}) = f^\lambda\,(b(\lambda') - b(\lambda)),$$

where $b(\mu)$ is defined by equation (7.103).

7.52. [2+] Let λ be a partition of N of rank r. For $1 \le i \le r$, let $\mu_i = h(i,i)$, the hook length of λ at (i,i). Set $\mu = (\mu_1, \mu_2, \ldots, \mu_r)$, so μ is also a partition of N. Show that

$$\chi^\lambda(\mu) = (-1)^t,$$

where $t = \sum_{i=1}^r(\lambda_i' - i)$. Moreover, if $\chi^\lambda(\nu) \neq 0$, then show that $\nu \le \mu$ (dominance order).

7.53. [2+] Let λ be a partition of N of rank r. Show that

$$\sum_w \chi^\lambda(w) = f^\lambda\,(-1)^{t(\lambda)}\prod_{i=1}^r(\lambda_i - 1)!(\lambda_i' - 1)!,$$

where w ranges over all permutations in $\mathfrak{S}_N$ with exactly r cycles, and where $t(\lambda) = \sum_{i=1}^r(\lambda_i' - i)$.

7.54. [3−] Prove the converse to Proposition 7.17.7, i.e., if $\lambda \vdash n$ and $\chi^\lambda(\mu) = 0$ whenever μ has an even part, then $n = \binom{m}{2}$ and $\lambda = (m-1, m-2, \ldots, 1)$ for some m.

7.55. **a.** [2+] Let $\rho^\lambda : \mathfrak{S}_n \to \mathrm{GL}(m, \mathbb{C})$ be an irreducible representation of $\mathfrak{S}_n$ with character χ^λ (so $m = f^\lambda$). Show that $\rho^\lambda(\mathfrak{S}_n) \subset \mathrm{SL}(m, \mathbb{C})$ if and only if

$$f^\lambda \left[\frac{b(\lambda') - b(\lambda)}{\binom{n}{2}} \right] \equiv f^\lambda \pmod 4, \tag{7.177}$$

where $b(\lambda)$ is defined by equation (7.103).

* **b.** [5−] Is there a simpler criterion? Is it possible to count the number of λ's satisfying (7.177)?

7.56. **a.** [2−] Given a skew shape θ, let θ^r denote the skew shape obtained by rotating θ 180°. For instance $(432/2)^r = 442/21$. Show that $s_\theta = s_{\theta^r}$.

b. [1+] Let $\alpha = (\alpha_1, \ldots, \alpha_k)$ be a composition, and let $\tilde{\alpha} = (\alpha_k, \ldots, \alpha_1)$. Show that $s_{B_\alpha} = s_{B_{\tilde{\alpha}}}$, where B_β denotes the border strip corresponding to β as defined in Section 7.23.

7.57. [2] Let $\lambda \vdash n$. How many border strips does λ have? In other words, how many partitions μ are there such that $\mu \subseteq \lambda$ and λ/μ is a (nonempty) border strip?

7.58. [2] Show that the number of odd hook lengths minus the number of even hook lengths of a partition λ is a triangular number.

7.59. This exercise deals with some basic combinatorial properties of border strips and hooks. Let λ be a partition, and let $p \in \mathbb{P}$. As noted in the solution to Exercise 7.57, there is a simple bijection between p-hooks (i.e., hooks of size p) of λ and border strips of λ of size p. Let D_λ denote the diagram of λ with its left-hand edge and upper edge extended to infinity, as shown in Figure 7-17 for $\lambda = (3,3,1)$. Put a 0 next to each vertical edge of the "lower envelope" of D_λ (whose definition should be clear from Figure 7-17), and a 1 next to each horizontal edge. If we read these numbers as we move north and east along the lower envelope, then we obtain an infinite binary sequence $C_\lambda = \cdots c_{-2}c_{-1}c_0c_1c_2\cdots$. For instance,

$$C_{331} = \cdots 0010110011 \cdots.$$

We regard a translate $\cdots b_{-1}b_0b_1 \cdots$ of C_λ, where $b_i = c_{m+i}$ for some fixed $m \in \mathbb{Z}$, as being the same as C_λ. Thus the choice of which term of C_λ is labeled c_0 is arbitrary. Clearly the map $\lambda \mapsto C_\lambda$ is a one-to-one correspondence between partitions and infinite binary sequences beginning with infinitely many 0's and ending with infinitely many 1's. The size $|\lambda|$ of λ is equal to the number of pairs $i < j$ with $c_i = 1$ and $c_j = 0$.

a. [2−] Show that there is a (natural) one-to-one correpondence between the p-hooks of λ and integers i such that $c_i = 1$ and $c_{i+p} = 0$.

b. [2] Show that removing a border strip of size p from λ is equivalent to choosing i with $c_i = 1$ and $c_{i+p} = 0$, and then replacing c_i with 0 and c_{i+p} with 1.

c. [2] Let θ be a border strip of λ of size p, and let $\lambda\backslash\theta$ denote the partition obtained by removing θ from λ. Show that $\lambda\backslash\theta$ has exactly one less hook length divisible by p than λ.

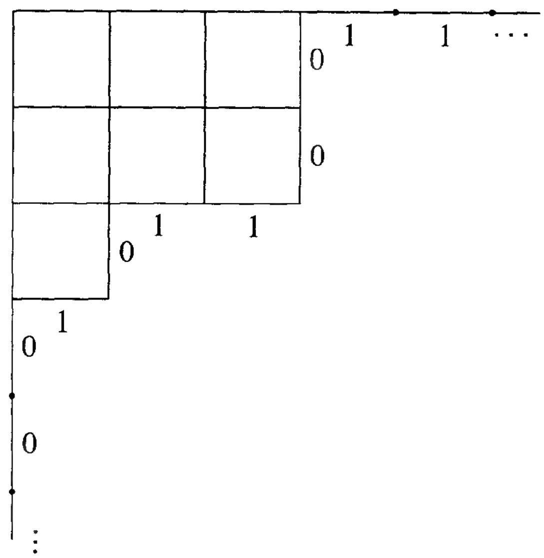


Figure 7-17. The coding C_λ of the partition $\lambda = (3,3,1)$.

d. [2+] Start with a partition λ, and continually remove border strips of size p until unable to do so. Show that the partition μ that remains is independent of the order in which the border strips are removed. The partition μ is called the *p-core* of λ, and a partition with no border strips (or hooks) of size p (equivalently, of size divisible by p) is called a *p-core*.

e. [2+] Let μ be a p-core. Let $Y_{p,\mu}$ be the set of all partitions whose p-core is μ. Define $\lambda \le \nu$ in $Y_{p,\mu}$ if λ can be obtained from ν by removing border strips of size p. Show that $Y_{p,\mu} \cong Y^k$, where Y denotes Young's lattice. Deduce that if $f_\mu(n)$ is the number of partitions of n with p-core μ, then

$$\sum_{n\ge 0} f_\mu(n)x^n = x^{|\mu|}\prod_{i\ge 1}\left(1-x^{pi}\right)^{-p}. \tag{7.178}$$

f. [2+] Let $n \in \mathbb{P}$. Show that the following three numbers are equal.

(i) The number of p-cores of size n.

(ii) The number of solutions $(x_1,\dots,x_{p-1}) \in \mathbb{N}^{p-1}$ to the equation

$$\sum_{i=1}^{p-1}\left[ix_i + p\binom{x_i}{2}\right] - \binom{x_1+\cdots+x_{p-1}}{2} = n.$$

(iii) The coefficient of x^n in

$$\prod_{i\ge 1}(1-x^{pi})^p(1-x^i)^{-1}.$$

g. [2] When $p=2$, find all partitions in (f)(i) explicitly. What identity results from the equality of (i) and (iii)?

h. [2] Let $C_p(n)$ be the set of all $\lambda \vdash pn$ whose p-core is empty. Let f_p^λ be the number of border strip tableaux τ of shape λ such that all the border

strips appearing in τ are of size p. Show that

$$\sum_{\lambda \in C_p(n)} \left(f_p^\lambda\right)^2 = p^n n!. \tag{7.179}$$

7.60. a. [2+] Let $\theta = \lambda/\mu$ be a border strip of size rs, where $r, s \in \mathbb{P}$. Show that there are partitions $\mu = \mu^0 \subset \mu^1 \subset \cdots \subset \mu^r = \lambda$ such that each skew shape μ^i/μ^{i-1} is a border strip of size s. Deduce in particular that if m is a hook length of λ and $k|m$, then k is also a hook length of λ.

b. [2+] Let $\lambda, \mu \vdash n$, with $\ell(\mu) = \ell$. Show that if $\chi^\lambda(\mu) \neq 0$, then the product $H_\lambda(q) := \prod_{u \in \lambda}(1 - q^{h(u)})$ is divisible (in $\mathbb{Z}[q]$) by $\prod_{i=1}^{\ell}(1 - q^{\mu_i})$. Here $h(u)$ denotes the hook length of λ at u.

7.61. [3−] Let $\lambda \vdash kn$. Show that

$$\left\langle h_n\left(x_1^k, x_2^k, \ldots\right), s_\lambda\right\rangle = 0 \text{ or } 1,$$

and give a rule for deciding which. In particular, show that this number is 0 unless λ has an empty k-core.

7.62. [2] Show that if $\lambda \vdash n$ and $\mu \vdash k \leq n$, then

$$\chi^\lambda(\mu 1^{n-k}) = \sum_{\nu \vdash k} f^{\lambda/\nu} \chi^\nu(\mu).$$

Here $\mu 1^{n-k}$ denotes the partition $\mu \cup \langle 1^{n-k} \rangle$.

7.63. a. [2+] For $\lambda \vdash n$ define

$$d_\lambda = \sum_{w \in \mathfrak{D}_n} \chi^\lambda(w),$$

where $\mathfrak{D}_n$ denotes the set of all derangements (permutations without fixed points) in $\mathfrak{S}_n$. Show that

$$\sum_{\lambda \vdash n} d_\lambda s_\lambda = \sum_{k=0}^{n} (-1)^{n-k} (n)_k h_1^{n-k} h_k.$$

b. [2+] Deduce from (a) that for $1 \leq k \leq n$,

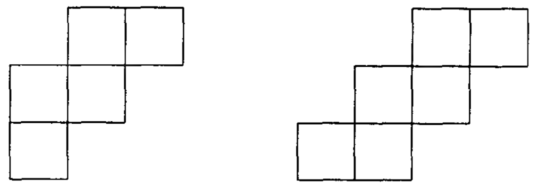
$$d_{\langle j, 1^{n-j} \rangle} = (-1)^{n-j} \binom{n}{j} D_j + (-1)^{n-1} \binom{n-1}{j},$$

where $D_j = \#\mathfrak{D}_j$ (discussed in Example 2.2.1).

7.64. a. [2] For a skew shape λ/μ where $|\lambda/\mu| = n$, define the skew character $\chi^{\lambda/\mu}$ of $\mathfrak{S}_n$ by $\operatorname{ch} \chi^{\lambda/\mu} = s_{\lambda/\mu}$, so $\deg \chi^{\lambda/\mu} = f^{\lambda/\mu}$. Now fix n and set $m = \lfloor \frac{1}{2}(n+2) \rfloor$. Define the skew shape

$$\tau_n = \begin{cases} (m, m-1, \ldots, 1)/(m-2, m-3, \ldots, 1), & n \text{ odd} \\ (m, m-1, \ldots, 2)/(m-2, m-3, \ldots, 1), & n \text{ even.} \end{cases}$$

Thus τ_n is a "staircase border strip," e.g.,



τ_5 τ_6

Let $E_n = \deg \chi^{\tau_n}$. Show that E_n is an Euler number, as defined at the end of Chapter 3.16, so

$$\sum_{n\geq 0} E_n \frac{x^n}{n!} = \tan x + \sec x.$$

b. [2+] Show that if $n = 2k+1$ and $\mu \vdash n$, then

$$\chi^{\tau_n}(\mu) = \begin{cases} 0 & \text{if } \mu \text{ has an even part} \\ (-1)^{k+r} E_{2r+1} & \text{if } \mu \text{ has } 2r+1 \text{ odd parts and no even parts.} \end{cases}$$

c. [2+] Show that if $n = 2k$ and $\mu \vdash n$, then

$$\chi^{\tau_n}(\mu) = (-1)^{k+r+e} E_{2r},$$

if μ has $2r$ odd parts and e even parts.

7.65. a. [2+] Let ψ_n be a character of $\mathfrak{S}_n$ for each $n \in \mathbb{P}$. Let us call the sequence $\psi_1, \psi_2, \ldots$ *elementary* if for all $w \in \mathfrak{S}_n$ we have that $\psi_n(w)$ is equal either to $\pm \deg \psi_m$ for some $m \leq n$ or to 0. For instance, the characters of the regular representations are elementary, as are the skew characters χ^{τ_n} of Exercise 7.64. Now define ψ_n by the condition that for $\lambda \vdash n$, $\langle \psi_n, \chi^\lambda \rangle$ is equal to the number of SYTs of shape λ whose largest descent has the same parity as n, where by convention every SYT has a descent at 0. For instance, $\psi_1 = 0$, $\psi_2 = \chi^2$, $\psi^3 = \chi^{21}$, $\psi^4 = \chi^4 + \chi^{31} + \chi^{22} + \chi^{211}$. Show that $\psi_1, \psi_2, \ldots$ is elementary, with $\deg \psi_n = D_n$, the number of derangements (permutations without fixed points) in $\mathfrak{S}_n$. Find $\psi_n(w)$ explicitly.

b. [5−] What other "interesting" elementary sequences are there? Can all elementary sequences be completely classified?

7.66. a. [3−] Let λ/μ be a skew shape. Define a *border strip decomposition* of λ/μ to be a partitioning of the squares of λ/μ into (nonempty) border strips. (We are not concerned with inserting the border strips in a particular order, as is the case for border strip tableaux.) For instance, Figure 7-18 shows a border strip decomposition of the shape 8877/211. Show that the number $d(\lambda/\mu)$ of border strip decompositions of λ/μ is a product of Fibonacci numbers. For instance, $d(8877/211) = 2 \cdot 3^2 \cdot 5 \cdot 8^2 \cdot 13 \cdot 21^2 \cdot 34$.

b. [3−] More generally, let

$$D_{\lambda/\mu}(q) = \sum_K q^{|\lambda/\mu| - \#K},$$

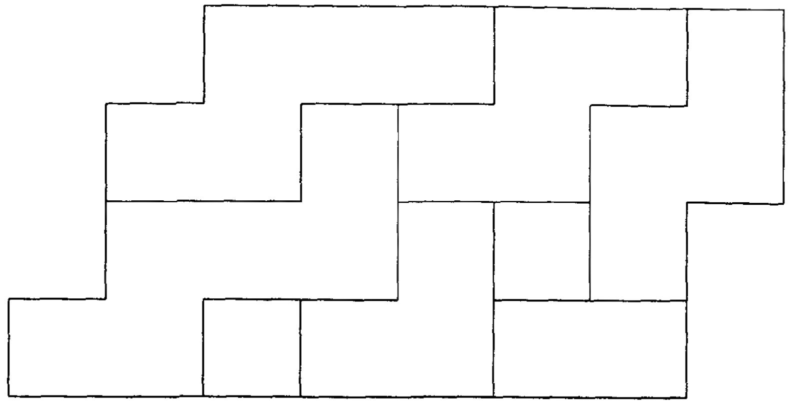


Figure 7-18. A border strip decomposition of the shape 8877/211.

where K ranges over all border strip decompositions of λ/μ and $\#K$ is the number of border strips appearing in K. Show that $D_{\lambda/\mu}(q)$ is a product of polynomials of the form $\sum_i \binom{m-i}{i} q^i$.

7.67. a. [2−] Let $0 \le s \le n-1$ and $\lambda \vdash n$. Show that if $w \in \mathfrak{S}_n$ is an n-cycle, then

$$\chi^\lambda(w) = \begin{cases} (-1)^s & \text{if } \lambda = \langle n-s, 1^s \rangle \\ 0 & \text{otherwise.} \end{cases}$$

b. [3] Let G be a finite group with conjugacy classes $C_1, \ldots, C_t$. Fix w in some class C_k, and let $i_1, \ldots, i_m \in [t]$. Let $\chi^1, \ldots, \chi^t$ be the irreducible characters of G, and set $d_r = \deg \chi^r$. Write χ_i^r for the common value of χ^r at any $v \in C_i$. Show that the number of m-tuples $(u_1, \ldots, u_m) \in G^m$ such that $u_j \in C_{i_j}$ and $u_1 \cdots u_m = w$ is equal to

$$\frac{\prod_{j=1}^m |C_{i_j}|}{|G|} \sum_{r=1}^t \frac{1}{d_r^{m-1}} \chi_{i_1}^r \cdots \chi_{i_m}^r \bar{\chi}_k^r. \tag{7.180}$$

c. [2] Fix $m \ge 1$. Use (a) and (b) to show that the number of m-tuples $(u_1, \ldots, u_m)$ of n-cycles $u_i \in \mathfrak{S}_n$ satisfying $u_1 u_2 \cdots u_m = 1$ is equal to

$$\frac{(n-1)!^{m-1}}{n} \sum_{i=0}^{n-1} (-1)^{im} \binom{n-1}{i}^{-(m-2)}. \tag{7.181}$$

d. [2+] When $m = 3$, show that the above sum is equal to 0 if n is even, and to $2(n-1)!^2/(n+1)$ if n is odd.

7.68. a. [3−] Let G be a finite group of order g. Given $w \in G$, let $f(w)$ be the number of pairs $(u, v) \in G \times G$ satisfying $w = uvu^{-1}v^{-1}$ (the *commutator* of u and v). Thus f is a class function on G and hence a linear combination $\sum c_\chi \chi$ of irreducible characters χ of G. Show that the multiplicity c_χ of χ in f is equal to $g/\chi(1)$. Since $\chi(1) \mid g$, it follows that f is a character of G.

b. [5] Find an explicit G-module M whose character is f. This would provide a new proof of the basic result that $\chi(1)$ divides the order of G.

c. [2] Deduce from (a) that

$$\frac{1}{n!}\sum_{u,v\in\mathfrak{S}_n} p_{\rho(uvu^{-1}v^{-1})} = \sum_{\lambda\vdash n} H_\lambda s_\lambda, \tag{7.182}$$

where H_λ denotes the product of the hook lengths of λ.

d. [2+] Let n be an odd positive integer. Show that the number f_n of ways to write the n-cycle $(1,2,\ldots,n)\in\mathfrak{S}_n$ in the form $uvu^{-1}v^{-1}$ $(u,v\in\mathfrak{S}_n)$ is equal to $2n\cdot n!/(n+1)$.

e. [1+] Let $\kappa(w)$ denote the number of cycles of a permutation $w\in\mathfrak{S}_n$. Deduce from (c) that

$$\frac{1}{n!}\sum_{u,v\in\mathfrak{S}_n} q^{\kappa(uvu^{-1}v^{-1})} = \sum_{\lambda\vdash n}\prod_{t\in\lambda}[q+c(t)],$$

where $c(t)$ denotes the content of the square t.

f. [3−] Show that if u, v are chosen at random (uniformly, independently) from $\mathfrak{S}_n$, then the expected number E_n of cycles of $uvu^{-1}v^{-1}$ is

$$E_n = H_n + \frac{1}{n!}\left(\sum_{\substack{1\le i\le n\\ i\text{ odd}}}\frac{i!\,(n-i)!}{n-i+1} + \frac{(-1)^n}{2}\sum_{i=0}^{\lfloor (n-1)/2\rfloor} i!^2\,(n-1-2i)!\right), \tag{7.183}$$

where H_n denotes the harmonic series $H_n = 1+\frac{1}{2}+\cdots+\frac{1}{n}$. Note that H_n is the expected number of cycles of a random permutation in $\mathfrak{S}_n$, so the remaining terms in (7.183) are a "correction."

g. [3−] Fix $j\in\mathbb{P}$. Show that if u, v are chosen at random (uniformly, independently) from $\mathfrak{S}_n$, then the expected number e_{nj} of j-cycles of $uvu^{-1}v^{-1}$ is given by

$$e_{nj} = \frac{1}{j}\left(1+\frac{1}{\binom{n}{j}}\sum_{i=0}^{j-1}{}' \frac{(-1)^i}{\binom{j-1}{i}}\frac{n-j+i+1}{n-2j+i+1}\right),$$

where $\sum'$ indicates that we are to omit the term $i=2j-n-1$ when $2j>n$.

7.69. a. [2+] Expand $\frac{1}{n!}\sum_{w\in\mathfrak{S}_n} p_{\rho(w^2)}$ in terms of Schur functions.

b. [2] Show that the column sums of the character table of $\mathfrak{S}_n$ are nonnegative. These are the numbers $\sum_{\lambda\vdash n}\chi^\lambda(w)$, where $w\in\mathfrak{S}_n$ is fixed. (For the row sums of the character table of a finite group, see Exercise 7.71(b).)

c. [3] Let k be a positive integer. Show that $\sum_{w\in\mathfrak{S}_n} p_{\rho(w^k)}$ is a *nonnegative* integer linear combination of Schur functions. Equivalently, the function $r_k = r_{n,k}:\mathfrak{S}_n\to\mathbb{Z}$ defined by

$$r_k(w) = \#\{u\in\mathfrak{S}_n\ :\ u^k = w\}$$

is a character of $\mathfrak{S}_n$.

d. [3−] Let G be a finite group, and let $f_1, \dots, f_m$ be class functions on G. Define a class function $F = F_{f_1,\dots,f_m}$ by

$$F(w) = \sum_{u_1 \cdots u_m = w} f_1(u_1) \cdots f_m(u_m).$$

Let χ be an irreducible character of G. Show that

$$\langle F, \chi \rangle = \left(\frac{|G|}{\chi(1)} \right)^{m-1} \langle f_1, \chi \rangle \cdots \langle f_m, \chi \rangle. \tag{7.184}$$

e. [2−] Show that equation (7.184) in the case $G = \mathfrak{S}_n$ and $m = 2$ is equivalent to the following result. Let $\tilde{s}_\lambda = H_\lambda s_\lambda$, called an *augmented Schur function*. Define a bilinear product $\Box$ on Λ^n by

$$p_\lambda \Box p_\mu = \frac{z_\lambda z_\mu}{n!} \sum_{\substack{\rho(u)=\lambda \\ \rho(v)=\mu}} p_{\rho(uv)} \qquad (\lambda, \mu \vdash n),$$

where the sum ranges over all $u, v \in \mathfrak{S}_n$ such that $\rho(u) = \lambda$ and $\rho(v) = \mu$. Then for $\lambda, \mu \vdash n$ we have

$$\tilde{s}_\lambda \Box \tilde{s}_\mu = \delta_{\lambda\mu} \tilde{s}_\lambda, \tag{7.185}$$

i.e., the augmented Schur functions are orthogonal idempotents with respect to $\Box$.

f. [1+] Let $(a_1, \dots, a_m) \in \mathbb{Z}^m$, and define a class function $h = h_{a_1,\dots,a_m}$ on $\mathfrak{S}_n$ by

$$h(w) = \#\left\{(u_1, \dots, u_m) \in \mathfrak{S}_n^m \,:\, w = u_1^{a_1} \cdots u_m^{a_m}\right\}.$$

Show that h is a character of $\mathfrak{S}_n$.

g. [2] Let G be any finite group, and let $w \in G$. Find the number of pairs $(u, v) \in G \times G$ satisfying $w = uvu^2vuv$.

h. [2−] Fix $w \in \mathfrak{S}_n$. Show that the number $f(w)$ of solutions $(u, v) \in \mathfrak{S}_n \times \mathfrak{S}_n$ to the equation $w = uvu^{-1}v^{-1}$ is equal to the number $g(w)$ of solutions $(u, v) \in \mathfrak{S}_n \times \mathfrak{S}_n$ to the equation $w = u^2v^2$. Give an algebraic and a bijective proof.

i. [2+] Let $\gamma = \gamma(x_1, \dots, x_r)$ be an element of the free group F_r on the generators $x_1, \dots, x_r$. If G is a finite group and $w \in G$, let

$$f_{\gamma,G}(w) = \#\{(u_1, \dots, u_r) \in G^r \,:\, \gamma(u_1, \dots, u_r) = w\}.$$

Write $x = x_1$ and $y = x_2$, and let $k \in \mathbb{P}$. Show that for $\gamma = xy^kxy^{-k}$ and $\gamma = xy^kx^{-1}y^{-k}$, the class functions $f_{\gamma,\mathfrak{S}_n}$ are characters of $\mathfrak{S}_n$ (for all $n \in \mathbb{P}$).

j. [3] Preserve the notation of (i). Suppose that all characters of G are integer-valued, in which case we say that G is an IC*-group*. (Equivalently, if two elements of G generate the same cyclic subgroup, then they are conjugate. See e.g. [**142**, Chap. 13.1, Cor. 2].) Show that for all $\gamma \in F_r$ and all finite IC-groups G, the class function $f_{\gamma,G}$ is a difference of two characters of G.

Even more strongly, when $r = 1$ show that $f_{\gamma,G}$ is a difference of two characters for *any* finite group G. Morever, for each conjugacy class C_i in

G, define the class function g_i by

$$g_i(w) = \begin{cases} |G|/|C_i|, & w \in C_i \\ 0 & \text{otherwise.} \end{cases}$$

If G is an IC-group, then it follows easily from the orthogonality properties of characters (see, e.g., [**142**, p. 20]) that g_i is a difference of characters. Show that if $r > 1$, then $f_{\gamma,G}$ is a $\mathbb{Z}$-linear combination of the g_i's. In particular, the symmetric function

$$\frac{1}{n!} \sum_{(u_1,\ldots,u_r)\in\mathfrak{S}_n^r} p_{\rho(\gamma(u_1,\ldots,u_r))}$$

is p-integral (and hence s-integral).

k. [5−] Preserve the notation of (i). For what γ is $f_{\gamma,\mathfrak{S}_n}$ a character of $\mathfrak{S}_n$ for all $n \geq 1$? For what γ is $f_{\gamma,G}$ a character of G for all finite groups G?

7.70. [3−] Let $x^{(1)}, \ldots, x^{(k)}$ be disjoint sets of variables, where $k \in \mathbb{N}$, and let H_λ denote the product of the hook lengths of λ. Show that

$$\sum_{\lambda\vdash n} H_\lambda^{k-2} s_\lambda\big(x^{(1)}\big)\cdots s_\lambda\big(x^{(k)}\big) = \frac{1}{n!} \sum_{\substack{w_1\cdots w_k=\mathrm{id} \\ \text{in}\,\mathfrak{S}_n}} p_{\rho(w_1)}\big(x^{(1)}\big)\cdots p_{\rho(w_k)}\big(x^{(k)}\big). \tag{7.186}$$

Note that the case $k = 2$ is just the Cauchy identity (in virtue of Proposition 7.7.4). What do the cases $k = 0$ and $k = 1$ say?

7.71. a. [2+] Show that the following two characters of a finite group G are the same:

(i) The character of the action of G on itself given by conjugation (in other words, the permutation representation $\rho : G \to \mathfrak{S}_G$ defined by $\rho(x)(y) = xyx^{-1}$, where $\mathfrak{S}_G$ is the group of permutations of G).

(ii) $\sum_\chi \chi\bar{\chi}$, where χ ranges over all irreducible characters of G.

b. [2+] Denote the above character by ψ_G, and let χ be an irreducible character of G. Show that

$$\langle\psi_G, \chi\rangle = \sum_K \chi(K),$$

where K ranges over all conjugacy classes of G and $\chi(K)$ denotes $\chi(w)$ for some $w \in K$. Thus $\langle\psi_G, \chi\rangle$ is the row sum of row χ of the character table of G. It is not *a priori* obvious that these row sums are nonnegative. (For the column sums of the character table of $\mathfrak{S}_n$, see Exercise 7.69(b).)

c. [2] Now let $G = \mathfrak{S}_n$, and write $\psi_n = \psi_G$. Show that $\mathrm{ch}\,\psi_n = \sum_{\lambda\vdash n} p_\lambda$, so $\sum_{n\geq 0} \mathrm{ch}\,\psi_n = \prod_{i\geq 1}(1 - p_i)^{-1}$.

d. [3] Show that $\kappa_\lambda := \langle\psi_n, \chi^\lambda\rangle > 0$, with the sole exception $n = 2$, $\lambda = (1, 1)$.

e. [5−] Is there a "nice" combinatorial interpretation of the numbers κ_λ?

7.72. [3−] Let V be a vector space over a field K of characteristic 0 ($\mathbb{Q}$ will do) with basis $v_1, \ldots, v_n$. $\mathfrak{S}_n$ acts on V by permuting coordinates, i.e., $w \cdot v_i = v_{w^{-1}(i)}$.

Hence $\mathfrak{S}_n$ acts on the k-th exterior power $\Lambda^k V$ in a natural way, viz.,

$$w(v_i \wedge v_j \wedge \cdots) = v_{w^{-1}(i)} \wedge v_{w^{-1}(j)} \wedge \cdots.$$

Show that the character of this action is equal to $\chi^{\lambda^{k-1}} + \chi^{\lambda^k}$, where $\lambda^j = \langle n-j, 1^j \rangle$. (Set $\chi^{\lambda^{-1}} = 0$.)

7.73. [3−] As in Exercise 7.72, $\mathfrak{S}_n$ also acts on the polynomial ring $K[v_1, \ldots, v_n]$ (= the symmetric algebra $S(V^*)$) in a natural way, viz.,

$$w\left(v_1^{a_1} v_2^{a_2} \cdots\right) = v_{w^{-1}(1)}^{a_1} v_{w^{-1}(2)}^{a_2} \cdots.$$

Let ψ^k denote the character of this action on the forms (homogeneous polynomials) of degree k, so $\deg \psi^k = \binom{n+k-1}{k}$. Show that

$$\sum_{k \geq 0} (\mathrm{ch}\, \psi^k) q^k = \sum_{\lambda \vdash n} s_\lambda(1, q, q^2, \ldots) s_\lambda,$$

where $s_\lambda(1, q, q^2, \ldots)$ is given explicitly by Corollary 7.21.3.

7.74. [3−] Let φ^λ be the irreducible representation of $\mathrm{GL}(n, \mathbb{C})$ with character $s_\lambda(x_1, \ldots, x_n)$, as explained in Appendix 2. We may regard $\mathfrak{S}_n$ as the subgroup of $\mathrm{GL}(n, \mathbb{C})$ consisting of the $n \times n$ permutation matrices. Thus φ^λ restricts to a representation of $\mathfrak{S}_n$; let ξ^λ denote its character. Show that for $\mu \vdash n$,

$$\langle \mathrm{ch}\, \xi^\lambda, s_\mu \rangle = \langle s_\lambda, s_\mu[h] \rangle, \tag{7.187}$$

where $s_\mu[h]$ denotes the plethysm of s_μ with the symmetric function $h = h_0 + h_1 + h_2 + \cdots$. Note that Exercise 7.72 corresponds to the case $\lambda = \langle 1^k \rangle$, while Exercise 7.73 corresponds to $\lambda = (k)$.

7.75. **a.** [2+] Fix positive integers n and k. Let M denote the multiset $\{1^n, 2^n, \ldots, k^n\}$. The action of $\mathfrak{S}_k$ on $[k]$ induces an action of $\mathfrak{S}_k$ on the set $\binom{M}{j}$ of j-element submultisets of M. Let $v_j(\lambda)$ denote the multiplicity of the irreducible character χ^λ (where $\lambda \vdash k$) in the character of this action. Show that

$$\sum_j v_j(\lambda) q^j = s_\lambda(1, q, \ldots, q^n).$$

b. [3−] Let $\mathbb{Q}S$ denote the $\mathbb{Q}$-vector space with basis S. Define a linear map $U_j : \mathbb{Q}\binom{M}{j} \to \mathbb{Q}\binom{M}{j+1}$ by

$$U_j(X) = \sum_{Y \supset X} Y,$$

where $X \in \binom{M}{j}$ and Y ranges over all elements of $\binom{M}{j+1}$ that contain X. Show that U_j commutes with the action of $\mathfrak{S}_k$, and that if $j < kn/2$ then U_j is injective.

c. [2+] Deduce that if $s_\lambda(1, q, \ldots, q^n) = \sum_{j=0}^{kn} a_j q^j$, then $a_j = a_{kn-j}$ (this is easy to do directly) and $a_0 \leq a_1 \leq \cdots \leq a_{\lfloor kn/2 \rfloor}$. In other words, the polynomial $s_\lambda(1, q, \ldots, q^n)$ is *symmetric* and *unimodal*.

d. [2−] Show that the q-binomial coefficient $\binom{n}{k}$ is symmetric and unimodal.

7.76. **a.** [2] The *rank* of a finite group G acting transitively on a set T is defined to be the number of orbits of G acting in the obvious way on $T \times T$, i.e., $w \cdot (s, t) = (w \cdot s, w \cdot t)$. Thus G is doubly transitive if and only if rank $G = 2$. Let χ be the character of the action of G on T. Show that $\langle \chi, \chi \rangle = \operatorname{rank} G$.

b. [2] Find the rank of the natural action of $\mathfrak{S}_n$ on the set $\mathfrak{S}_n / \mathfrak{S}_\alpha$ of left cosets of the Young subgroup $\mathfrak{S}_\alpha$.

c. [2+] Give a direct bijective proof of (b).

7.77. **a.** [3−] If H and K are subgroups of a group G, then a *double coset* of (H, K) is a set $HwK = \{uwv | u \in H, v \in K\}$ for a fixed $w \in G$. The distinct double cosets of (H, K) partition G into pairwise disjoint nonempty subsets (not necessarily of the same cardinality). Show that when G is finite, the number of double cosets of (H, K) is given by

$$\left\langle \operatorname{ind}_H^G 1_H, \operatorname{ind}_K^G 1_K \right\rangle.$$

b. [2+] For any G (not necessarily finite), show that the number of double cosets of (H, H) is equal to the rank (as defined in Exercise 7.76) of G acting on G/H by left multiplication.

c. [2+] Let $G = \mathfrak{S}_n$, and let H and K be Young subgroups, say $H = \mathfrak{S}_\alpha$ and $K = \mathfrak{S}_\beta$. Interpret the number of double cosets of (H, K) in a simple combinatorial way, and give a combinatorial proof.

7.78. Let f and g be class functions (over $\mathbb{Z}$ or a field of characteristic 0) on a finite group G. Define the *Kronecker* (or *tensor*) product fg by $fg(w) = f(w)g(w)$, so fg is also a class function. Given (finite-dimensional) representations $\varphi : G \to \mathrm{GL}(V)$ and $\rho : G \to \mathrm{GL}(W)$, then define the tensor product representation $\varphi \otimes \rho : G \to \mathrm{GL}(V \otimes W)$ by

$$w \cdot (x \otimes y) = w \cdot x \otimes w \cdot y \qquad \text{(diagonal action)}.$$

Let χ_φ and χ_ρ denote the characters of φ and ρ, respectively. Then the character $\chi_{\varphi \otimes \rho}$ of $\varphi \otimes \rho$ is just the Kronecker product $\chi_\varphi \chi_\rho$, so $\chi_\varphi \chi_\rho$ is a nonnegative integer linear combination of irreducible characters. In particular, for $G = \mathfrak{S}_n$ and $\lambda, \mu \vdash n$, we have

$$\chi^\lambda \chi^\mu = \sum_{\nu \vdash n} g_{\lambda\mu\nu} \chi^\nu, \tag{7.188}$$

for certain *nonnegative* integers $g_{\lambda\mu\nu}$. Define the *internal product* $s_\lambda * s_\mu$ by

$$s_\lambda * s_\mu = \sum_\nu g_{\lambda\mu\nu} s_\nu,$$

and extend to all of Λ by bilinearity. Clearly $*$ is associative and commutative, and $h_n * f = f$ for $f \in \Lambda^n$.

a. [2−] Show that $g_{\lambda\mu\nu}$ is invariant under permuting the indices λ, μ, ν.

b. [2−] Show that

$$s_\lambda * s_\mu = \operatorname{ch} \chi^\lambda \chi^\mu := \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} \chi^\lambda(w) \chi^\mu(w) p_{\rho(w)}.$$

c. [2] Show that if $f \in \Lambda^n$ then $e_n * f = \omega f$.

d. [2] Show that $p_\lambda * p_\mu = z_\lambda \, p_\lambda \, \delta_{\lambda\mu}$.

e. [2] Let $\Lambda(x) \otimes \Lambda(y)$ be defined as at the end of Section 7.15, with the scalar product such that the elements $s_\mu(x)s_\nu(y)$ form an orthonormal basis. Let xy denote the set of variables $x_i y_j$. Show that for any $f, g, h \in \Lambda$ we have

$$\langle f(xy), g(x)h(y) \rangle = \langle f, g * h \rangle,$$

where the first scalar product takes place in $\Lambda(x) \otimes \Lambda(y)$ and the second in Λ. This gives a "basis-free" definition of the internal product, analogous to equation (7.67) for the ordinary product.

f. [2+] Show that

$$\prod_{i,j,k}(1 - x_i y_j z_k)^{-1} = \sum_{\lambda,\mu} s_\lambda * s_\mu(x) s_\lambda(y) s_\mu(z)$$
$$= \sum_{\lambda,\mu,\nu} g_{\lambda\mu\nu} s_\lambda(x) s_\mu(y) s_\nu(z).$$

g. [2+] More generally, show that if $x^{(1)}, \ldots, x^{(k)}$ are disjoint sets of variables, then

$$\prod_{i_1,\ldots,i_k} \left(1 - x^{(1)}_{i_1} \cdots x^{(k)}_{i_k}\right)^{-1}$$
$$= \sum_{\lambda^1,\ldots,\lambda^k \vdash n} \left\langle 1_{\mathfrak{S}_n}, \chi^{\lambda^1} \cdots \chi^{\lambda^k} \right\rangle s_{\lambda^1}\left(x^{(1)}\right) \cdots s_{\lambda^k}\left(x^{(k)}\right).$$

7.79. **a.** [3–] Show that if $\langle s_\lambda, s_\mu * s_\nu \rangle \neq 0$, then $\ell(\lambda) \leq \ell(\mu)\ell(\nu)$.

b. [3–] Suppose that $\ell(\lambda) \leq ab$. Show that there exist partitions μ, ν satisfying $\ell(\mu) \leq a$, $\ell(\nu) \leq b$, and $\langle s_\lambda, s_\mu * s_\nu \rangle \neq 0$.

c. [3] Prove the following strengthening of (a) and (b): for fixed $\mu, \nu \vdash n$, we have

$$\max\{\ell(\lambda) : \langle s_\lambda, s_\mu * s_\nu \rangle \neq 0\} = |\mu \cap \nu'|,$$

where $\mu \cap \nu'$ is obtained by intersecting the diagrams of μ and ν'. Dually, we have

$$\max\{\lambda_1 : \langle s_\lambda, s_\mu * s_\nu \rangle \neq 0\} = |\mu \cap \nu|.$$

7.80. **a.** [3–] Let $\lambda, \mu, \nu \vdash n$ satisfy $s_\lambda * s_\mu = a s_\nu$ for some $a \in \mathbb{P}$. Show that one of λ or μ is equal to (n) or $\langle 1^n \rangle$, and that $a = 1$.

b. [3] Let $\lambda, \mu, \nu, \sigma \vdash n$, where $\nu \neq \sigma$, satisfy $s_\lambda * s_\mu = a s_\nu + b s_\sigma$ for some $a, b \in \mathbb{P}$. Show that one of λ or μ has nontrivial (i.e., not (n) or $\langle 1^n \rangle$) rectangular shape, and the other is equal to $(n-1, 1)$ or $\langle 2, 1^{n-2} \rangle$, and that $a = b = 1$.

7.81. [2+] Show that for $\lambda \vdash n$,

$$s_\lambda * s_{n-1,1} = s_1 s_{\lambda/1} - s_\lambda.$$

7.82. **a.** [2] Show that

$$\sum_{\lambda \in \mathrm{Par}} s_\lambda * s_\lambda = \frac{1}{\prod_{i \geq 1}(1 - p_i)},$$

where $*$ denotes internal product.

b. [3−] Show that

$$\frac{\partial}{\partial p_1}\sum_{\ell(\lambda)\leq 2} s_\lambda * s_\lambda = \left(\sum_{n\geq 0} s_n\right)\left(\sum_{\ell(\lambda)\leq 3} s_\lambda\right).$$

7.83. a. [2+] Let χ and ψ be irreducible characters of a finite group G. Show that $\chi\psi$ is contained in the regular representation, i.e., $\langle\chi\psi,\phi\rangle\leq\phi(1)$ for any irreducible character ϕ of G.

b. [1+] Deduce from (a) that if $\lambda,\mu,\nu\vdash n$, then $\langle s_\lambda * s_\mu, s_\nu\rangle\leq f^\nu$ (the number of SYTs of shape ν).

7.84. a. [2+] Let $\lambda,\mu\vdash n$, with $\ell(\lambda)=\ell$. Show that

$$h_\lambda * s_\mu = \sum\prod_{i\geq 1} s_{\mu^i/\mu^{i-1}},$$

summed over all sequences $(\mu^0,\mu^1,\ldots,\mu^\ell)$ of partitions such that $\emptyset = \mu^0\subset\mu^1\subset\cdots\subset\mu^\ell=\mu$ and $|\mu^i/\mu^{i-1}|=\lambda_i$ for all $i\geq 1$.

b. [2+] Let $\lambda,\mu\vdash n$. Show that

$$h_\lambda * h_\mu = \sum_A \prod_{i,j=1}^{n} h_{a_{ij}},$$

where A ranges over all $n\times n$ $\mathbb{N}$-matrices (a_{ij}) with $\mathrm{row}(A)=\lambda$ and $\mathrm{col}(A)=\mu$.

7.85. [3] Fix $n\geq 1$. Given $\alpha=(\alpha_1,\ldots,\alpha_k)\in\mathbb{P}^k$ with $\alpha_1+\cdots+\alpha_k=n$, let $B(\alpha)$ denote the border strip whose ith row has α_i squares. Let $S=\{\alpha_1,\alpha_1+\alpha_2,\ldots,\alpha_1+\alpha_2+\cdots+\alpha_{k-1}\}\subseteq[n-1]$ and define $s_S=s_{B(\alpha)}$, the skew Schur function of shape $B(\alpha)$. Now let $S,T\subseteq[n-1]$ and $\lambda\vdash n$, and let $*$ denote internal product. Show that $\langle s_S * s_T, s_\lambda\rangle$ is equal to the number of triples $u,v,w\in\mathfrak{S}_n$ such that $uvw=1$, $D(u)=S$, $D(v)=T$, and if w is inserted into λ from right to left and from bottom to top, an SYT results. Note that the hook shapes $\langle n-k,1^k\rangle$ are border strips, so we have a combinatorial interpretation of the coefficients $g_{\lambda\mu\nu}$ when μ and ν are hooks.

Example. Let $n=3$ and $S=T=\{1\}$, so $s_S=s_T=s_{21}$. There are four triples $u,v,w\in\mathfrak{S}_3$ such that $uvw=1$ and $D(u)=D(v)=\{1\}$. In only one of these can we get an SYT by inserting w as required, viz., $u=312$, $v=213$, $w=321$. We can insert w exactly once into each $\lambda\vdash 3$, viz.,

$$\begin{array}{lll} 1\,2\,3 & 1\,2 & 1 \\ & 3 & 2 \\ & & 3 \end{array}.$$

Hence $s_{21} * s_{21} = s_3 + s_{21} + s_{111}$.

7.86. a. [3−] Given $\lambda,\mu\vdash n$, define

$$G_{\lambda\mu}(q) = s_\lambda * s_\mu(1,q,q^2,\ldots).$$

Show that $G_{\lambda\mu}(q)=P_{\lambda\mu}(q)H_\lambda(q)^{-1}$, where $P_{\lambda\mu}\in\mathbb{Z}[q]$ and $H_\lambda(q)$ is defined in Exercise 7.60(b).

b. [2+] Show that $P_{\lambda\mu}(1)=f^\mu$, the number of SYTs of shape μ.

c. [3] Show that if $\mu = \langle n-k, 1^k\rangle$, then $P_{\lambda\mu}(q)$ is the coefficient of t^k in the product

$$\prod_{\substack{(i,j)\in\lambda \\ (i,j)\neq(1,1)}} (q^{i-1} + tq^{j-1}).$$

d. [5] Show that the coefficients of $P_{\lambda\mu}(q)$ are nonnegative. This has been checked for $n \leq 9$. Note that (c) shows that the coefficients of $P_{\lambda\mu}$ are indeed nonnegative when μ is a hook.

7.87. [3−] Let x, y, z be three sets of variables. Show that

$$\prod_{i,j}\prod_{r\geq 1}\prod_{a_1,\ldots,a_r} \left(1 - x_i y_j z_{a_1}\cdots z_{a_r}\right)^{-1}$$

$$= \left[\prod_{k\geq 1}(1 - p_k(z))\right] \sum_{\substack{\lambda,\mu,\nu \\ |\lambda|=|\mu|}} s_\lambda * s_\mu(z) s_{\lambda/\nu}(x) s_{\mu/\nu}(y).$$

Here $a_1, \ldots, a_r$ range independently over the positive integers and $p_k(z) = \sum z_i^k$.

7.88. **a.** [3−] Let C_n be the cyclic subgroup of $\mathfrak{S}_n$ of order n generated by an n-cycle w. Let χ be the character of C_n defined by $\chi(w) = e^{2\pi i/n}$. Let $\psi_m = \psi_{m,n}$ denote the induction of χ^m to $\mathfrak{S}_n$, for $m \in \mathbb{Z}$. Show that

$$\operatorname{ch}\psi_m = \frac{1}{n}\sum_{d|n} \frac{\phi(d)}{\phi(d/(m,d))}\mu(d/(m,d)) p_d^{n/d}, \qquad (7.189)$$

where ϕ denotes Euler's totient function and (m,d) denotes the greatest common divisor of m and d.

b. [3−] Show that $\langle \psi_m, s_\lambda\rangle$ is equal to the number of SYTs T of shape λ satisfying $\operatorname{maj}(T) \equiv m \pmod{n}$.

c. [2−] Deduce from (a) and (b) that the number of SYTs T of shape λ satisfying $\operatorname{maj}(T) \equiv m \pmod{n}$ depends only on λ and $\gcd(m,n)$. Is there a bijective proof?

d. [3+] Let $y_k(\lambda)$ denote the number of SYTs T of shape λ satisfying $\operatorname{maj}(T) \equiv 1 \pmod{k}$. Show that if $\lambda \vdash n$ then

$$y_{n-1}(\lambda) \geq y_n(\lambda).$$

e. [2+] Let

$$J = \omega\sum_{n\geq 1}(-1)^{n-1}\operatorname{ch}(\psi_{1,n})$$

$$= \sum_{d\geq 1}\frac{\mu(d)}{d}\log(1 + p_d),$$

and let $H := 1 + h_1 + h_2 + \cdots$. Show that $J[H-1] = (H-1)[J] = h_1$, where brackets denote plethysm. In other words, J and $H-1$ are plethystic inverses of one another.

7.89. **a.** [3−] Let $a < b < c < \cdots$ be an ordered alphabet. A *Lyndon word* is a word $w_1 w_2 \cdots w_n$ in the alphabet which is lexicographically strictly less than all its nonidentity cyclic shifts. Thus $aabcaabbc$ is not a Lyndon word, since its cyclic shift $aabbcaabc$ is lexicographically smaller; nor is $abab$ a Lyndon word, since it is equal to its cyclic shift of length two. Let $f(\alpha)$ be the number of Lyndon words with α_1 a's, α_2 b's, etc., where $\alpha = (\alpha_1, \alpha_2, \ldots)$. Define

$$L_n(x) = \sum_\alpha f(\alpha)x^\alpha, \tag{7.190}$$

where α ranges over all weak compositions of n. For instance, $L_3 = m_{21} + 2m_{111}$. Show that

$$L_n = \frac{1}{n}\sum_{d|n} \mu(d) p_d^{n/d}, \tag{7.191}$$

where μ denotes the classical Möbius function of number theory.

b. [1+] Show that $L_n = \mathrm{ch}(\psi)$, where ψ denotes the character of $\mathfrak{S}_n$ obtained by inducing from a cyclic subgroup C_n generated by an n-cycle w to $\mathfrak{S}_n$ the character χ defined by $\chi(w) = e^{2\pi i/n}$. Deduce that $\langle L_n, s_\lambda\rangle \in \mathbb{N}$ for every $\lambda \vdash n$.

c. [1+] Even more strongly, show that $\langle L_n, s_\lambda\rangle$ is the number of standard Young tableaux T of shape λ satisfying $\mathrm{maj}(T) \equiv 1 \pmod n$.

d. [3−] Show that every word w in the letters $a, b, \ldots$ can be factored *uniquely* into a weakly decreasing (in lexicographic order) product of Lyndon words. For example, $bccbbcbaccaccabaabaa$ has the factorization $bcc \cdot bbc \cdot b \cdot acc \cdot acc \cdot ab \cdot aab \cdot a \cdot a$.

e. [2+] Given a word w as above, define its *Lyndon type* $\tau(w)$ to be the partition whose parts are the lengths of the Lyndon words in the factorization of w into a weakly decreasing product of Lyndon words. For instance, $\tau(dbca) = (2, 1, 1)$. Show that

$$\sum_w p_{\tau(w)} = n!h_n,$$

where w ranges over all permutations of n ordered letters. In other words, the distribution by Lyndon type of the permutations of an (ordered) n-set coincides with the distribution by cycle length.

f. [3−] Let M be a finite multiset on the set $\{a, b, \ldots\}$. Define

$$t_M = \sum_w p_{\tau(w)},$$

where w ranges over all permutations of M. Thus by (e), t_M is a multiset analogue of the cycle indicator of $\mathfrak{S}_n$. Define $L_{(i^m)}$ to be the plethysm $h_m[L_i]$ (as defined in Appendix 2). If $\lambda = \langle 1^{m_1} 2^{m_2} \cdots\rangle$, then let $L_\lambda = L_{(1^{m_1})} L_{(2^{m_2})} \cdots$. If $M = \{a^{r_1}, b^{r_2}, \ldots\}$ and μ is the partition with parts $r_1, r_2, \ldots$, then show that $t_M(y)$ is the coefficient of $m_\mu(x)$ in $\sum_\lambda L_\lambda(x)p_\lambda(y)$.

g. [3−] Show that

$$\sum_\lambda L_\lambda(x)p_\lambda(y) = \sum_\lambda p_\lambda(x)L_\lambda(y).$$

h. [2+] Deduce that t_M is s-positive (and s-integral).

i. [5−] Is there a "nice" combinatorial interpretation of the coefficients $\langle t_M, s_\lambda\rangle$?

7.90. **a.** [2] Let $\alpha = (\alpha_1, \ldots, \alpha_k)$ be a sequence of positive integers summing to n and let $|\lambda/\mu| = n$. Let $S = \{\alpha_1, \alpha_1 + \alpha_2, \ldots, \alpha_1 + \cdots + \alpha_{k-1}\}$. Show that the number of (skew) SYTs τ of shape λ/μ satisfying $D(\tau) \subseteq S$ (where $D(\tau)$ denotes the descent set of τ) is equal to the Kostka number $K_{\lambda/\mu,\alpha}$. Give a simple bijective proof.

b. [2] Use (a) to give a simple direct proof of Proposition 7.19.9.

7.91. Let $F(t) = \sum_{j\geq 0} f_j t^j$ be a formal power series, where $f_0 = 1$. Expand the product $F(t_1)F(t_2)\cdots$ as a linear combination of Schur functions $s_\lambda(t_1, t_2, \ldots)$. The coefficient of $s_\lambda(t_1, t_2, \ldots)$ is called (in the terminology of D. E. Littlewood [**88**, pp. 99–100 and Chap. VII]) the *Schur function* (indexed by λ) *of the series* F, and we will denote it by s_λ^F. Equivalently, if R is a (commutative) ring containing $f_1, f_2, \ldots$ and $\varphi : \Lambda \to R$ is the homomorphism defined by $\varphi(h_j) = f_j$, then $s_\lambda^F = \varphi(s_\lambda)$. Extend the definition of s_λ^F by defining $u^F = \varphi(u)$ for any $u \in \Lambda_R$.

a. [1] Show that if $F(t) = \prod_{i\geq 1}(1 - x_i t)^{-1}$, then $s_\lambda^F = s_\lambda(x)$. What if $F(t) = \prod_{i\geq 1}(1 + x_i t)$?

b. [1] Show that if $F(t) = \prod_{i=1}^{n}(1 - q^{i-1}t)^{-1}$, then

$$s_\lambda^F = q^{b(\lambda)} \prod_{u\in\lambda} \frac{1 - q^{n+c(u)}}{1 - q^{h(u)}},$$

where $b(\lambda)$, $c(u)$, and $h(u)$ have the same meaning as in Theorem 7.21.2.

c. [2+] Deduce from (b) that if

$$F(t) = \prod_{i\geq 0} \frac{1 - yq^i t}{1 - zq^i t},$$

then

$$s_\lambda^F = q^{b(\lambda)} \prod_{u\in\lambda} \frac{y - zq^{c(u)}}{1 - q^{h(u)}}. \tag{7.192}$$

d. [2−] Show that in general if $F(t) = \sum_{j\geq 0} f_j t^j$ and $\ell(\lambda) = \ell$, then

$$s_\lambda^F = \det\left(f_{\lambda_i - i + j}\right)_{i,j=1}^{\ell}.$$

e. [3+] Suppose that $F(t)$ is a nonconstant *polynomial* with complex coefficients (with $F(0) = 1$ as usual), so that s_λ^F is just a complex number. Show that the following four conditions are equivalent.

(i) Every zero of $F(t)$ is a negative real number.

(ii) For all partitions λ, s_λ^F is a nonnegative real number. Equivalently, when the product $F(t_1)F(t_2)\cdots$ is expanded as a linear combination of Schur functions $s_\lambda(t_1, t_2, \ldots)$, all the coefficients are nonnegative real numbers. In other words, $F(t_1)F(t_2)\cdots$ is s-positive.

(iii) When the product $F(t_1)F(t_2)\cdots$ is expanded as a linear combination of elementary symmetric functions $e_\lambda(t_1, t_2, \ldots)$, all the coefficients are nonnegative real numbers. In other words, $F(t_1)F(t_2)\cdots$ is e-positive. Equivalently, m_λ^F is a nonnegative real number for all partitions λ.

(iv) All coefficients of $F(t)$ are nonnegative real numbers, and the matrix $A = (p_{i+j}^F)_{i,j=0}^{n-1}$ is positive semidefinite. Here we set $p_0^F = \deg F$.

7.92. **a.** [3+] Let $A = (a_{ij})$ be an $n \times n$ real matrix such that every minor (= determinant of a square submatrix) is nonnegative. Define the symmetric function

$$F_A = \sum_{w \in \mathfrak{S}_n} a_{1,w(1)} a_{2,w(2)} \cdots a_{n,w(n)} p_{\rho(w)}.$$

Show that F_A is s-positive.

b. [5] Show that F_A is h-positive.

7.93. [2+] Let $u = u_1 \cdots u_m \in \mathfrak{S}_m$ and $v = v_1 \cdots v_n \in \mathfrak{S}_{[m+1,m+n]}$. Let $\mathrm{sh}(u, v)$ denote the set of *shuffles* of the words $u_1 \cdots u_m$ and $v_1 \cdots v_n$, i.e., $\mathrm{sh}(u, v)$ consists of all permutations $w_1 \cdots w_{m+n}$ of $[m+n]$ such that $u_1 \cdots u_m$ and $v_1 \cdots v_n$ are subsequences of w. Hence in particular $\#\mathrm{sh}(u, v) = \binom{m+n}{m}$. Let $\alpha = \mathrm{co}(u)$ and $\beta = \mathrm{co}(v)$, as defined at the beginning of Section 7.19. Show that

$$L_\alpha L_\beta = \sum_{w \in \mathrm{sh}(u,v)} L_{\mathrm{co}(w)}.$$

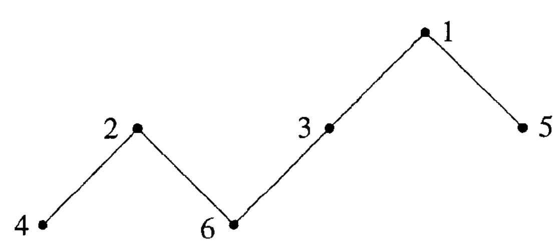
7.94. **a.** [2+] Let $\mathcal{Q}$ denote the ring of quasisymmetric functions (over $\mathbb{Q}$), and define a linear involution $\hat{\omega} : \mathcal{Q} \to \mathcal{Q}$ by $\hat{\omega}(L_\alpha) = L_{\bar{\alpha}}$, where if $\alpha \in \mathrm{Comp}(n)$ then $S_{\bar{\alpha}} = [n-1] - S_\alpha$. Show that $\hat{\omega}$ is an automorphism of $\mathcal{Q}$, and that $\hat{\omega}$ restricted to Λ coincides with the involution ω.

b. [3−] Let P be a finite graded poset of rank n with $\hat{0}$ and $\hat{1}$. Let $f \in I(P)$ (the incidence algebra of P) satisfy $f(t,t) = 1$ for all $t \in P$. Define

$$F_f = \sum_{\hat{0}=t_0 \le t_1 \le \cdots \le t_{k-1} < t_k = \hat{1}} f(t_0, t_1) f(t_1, t_2) \cdots f(t_{k-1}, t_k) \times x_1^{\rho(t_0,t_1)} x_2^{\rho(t_1,t_2)} \cdots x_k^{\rho(t_{k-1},t_k)},$$

using the same notation as equation (7.176). Clearly $F_f \in \mathcal{Q}^n$. Show that $\hat{\omega}(F_f) = (-1)^n F_{f^{-1}}$. Note that Exercise 7.48(b) is the special case $f = \zeta$.

7.95. **a.** [2] Given $S \subseteq [n-1]$, let $\alpha = \mathrm{co}(S)$ be the corresponding composition of n, as defined at the beginning of Section 7.19. Let B_α be the border strip whose i-th row from the bottom has length α_i, and write P_S as short for the poset P_{B_α} (where $P_{\lambda/\mu}$ is defined after Corollary 7.19.5). Given $w = w_1 w_2 \cdots w_n \in \mathfrak{S}_n$, let ω_w be the labeling of P_S obtained by inserting the numbers w_i into the squares of B_α from bottom to top and from left to right. For instance, if $\alpha = (2, 3, 1)$ (so $S = \{2, 5\}$) and $w = 426315$, then (P_S, ω_w) looks as follows:



Show that the Jordan–Hölder set $\mathcal{L}(P_S, \omega_w)$ consists of all permutations $v \in \mathfrak{S}_n$ such that $D(wv^{-1}) = S$.

b. [2+] Deduce from (a) the following statement. Given $w \in \mathfrak{S}_n$ and $S, T \subseteq [n-1]$, define

$$f(w, S, T) = \#\{(u, v) \in \mathfrak{S}_n \times \mathfrak{S}_n \ : \ uv = w, \ D(u) = S, \ D(v) = T\}.$$

Then $f(w, S, T) = f(w', S, T)$ whenever $D(w) = D(w')$.

7.96. [3] For $w \in \mathfrak{S}_n$ let $L_{\mathrm{co}(w)}$ denote the quasisymmetric function given by equation (7.89), where $\mathrm{co}(w)$ is defined at the beginning of Section 7.19. Define

$$T_n = \sum_{w \in \mathfrak{S}_n} L_{\mathrm{co}(w)} w,$$

regarded as an element of the group algebra $\mathcal{Q}\mathfrak{S}_n$ with coefficients in the ring $\mathcal{Q}$ of quasisymmetric functions. Thus T_n acts on $\mathcal{Q}\mathfrak{S}_n$ by left multiplication. Show that the eigenvalues of T_n are the power sums p_λ with multiplicity $n!/z_\lambda$, the number of permutations $w \in \mathfrak{S}_n$ of cycle type λ. What are the eigenvectors?

7.97. **a.** [2] Fix r, c, and $\lambda = (\lambda_1, \lambda_2, \ldots) \vdash t$. Let $f(n)$ be the number of plane partitions $\pi = (\pi_{ij})$ of n with main diagonal $(\pi_{11}, \pi_{22}, \ldots) = \lambda$ and with at most r rows and at most c columns. Set $F(x) = \sum_{n \geq 0} f(n) x^n$, and show that

$$F(x) = x^{-t} s_\lambda(x, x^2, \ldots, x^r) s_\lambda(x, x^2, \ldots, x^c).$$

b. [2+] Show that if $g(n)$ is the number of *symmetric* plane partitions with main diagonal λ, then

$$\sum_{n \geq 0} g(n) x^n = s_\lambda(x, x^3, x^5, \ldots).$$

7.98. **a.** [2+] Given $\lambda \vdash n$ and $(i, j) \in \lambda$, define

$$w(i, j) = \prod_{(k,l)} x_{k-l},$$

where the x_m's are indeterminates and the product ranges over all squares $(k, l) \in \lambda$ in the hook of (i, j), i.e., such that either $k = i$ and $l \geq j$, or $l = j$ and $k \geq i$. If now $\pi = (\pi_{ij})$ is a reverse plane partition of shape λ (allowing 0 as a part), then define

$$W(\pi) = \prod_{(i,j) \in \lambda} x_{j-i}^{\pi_{ij}}.$$

Show that

$$\sum_{\pi} W(\pi) = \prod_{u \in \lambda} [1 - w(u)]^{-1},$$

where π ranges over all reverse plane partitions of shape λ (allowing 0 as a part).

b. [3−] State and prove an analogous result for *symmetric* reverse plane partitions of shape λ (where $\lambda = \lambda'$).

7.99. [2+] Let $K_t(n)$ be the number of plane partitions of n with trace t. Show that if $0 \leq n \leq t$, then $K_t(n+t)$ is equal to the coefficient of x^n in the expansion

$$\prod_{i\geq 1}(1-x^i)^{-i-1} = 1 + 2x + 6x^2 + 14x^3 + 33x^4 + 70x^5 + 149x^6 + \cdots.$$

7.100. **a.** [3−] Let A and B be two $\mathbb{N}$-matrices with the same support, i.e., $a_{ij} \neq 0$ if and only $b_{ij} \neq 0$. If $A \xrightarrow{\text{RSK}} (P, Q)$ and $B \xrightarrow{\text{RSK}} (P', Q')$, then show that P and P' have the same first columns, and that Q and Q' have the same first columns.

b. [2+] Let $t_\lambda(n)$ denote the number of plane partitions $\pi = (\pi_{ij})$ whose shape is contained in λ and that satisfy $n = \operatorname{tr}(\pi) := \pi_{11} + \pi_{22} + \cdots$. Show that $t_\lambda(n)$ is a polynomial function of n of degree $|\lambda| - 1$.

c. [2+] Show that if λ is an $a \times b$ rectangle (i.e., λ has a parts, all equal to b), then

$$t_\lambda(n) = \binom{ab+n-1}{ab-1}.$$

7.101. **a.** [3−] Let δ_n be the staircase shape $(n-1, n-2, \ldots, 1)$, and let $f_n(m)$ denote the number of plane partitions, allowing 0 as a part, of shape δ_n and with largest part at most m. For instance, it follows from Exercise 6.19(vv) that $f_n(1) = C_n$ (a Catalan number). Show that

$$f_n(m) = \prod_{i=1}^{n-1} \frac{m+i}{i}\left(\prod_{j=2}^{i} \frac{2m+i+j-1}{i+j-1}\right) = \prod_{1\leq i<j\leq n} \frac{2m+i+j-1}{i+j-1}. \tag{7.193}$$

* **b.** [3−] More generally, let $g_{Md\ell}(m)$ denote the number of plane partitions, allowing 0 as a part, of shape $\lambda = (M-d, M-2d, \ldots, M-\ell d)$. Show that

$$g_{Md\ell}(m) = \prod_{\substack{u=(i,j)\in\lambda \\ \ell+c(u)\leq\lambda_i}} \frac{m+\ell+c(u)}{\ell+c(u)} \cdot \prod_{\substack{u=(i,j)\in\lambda \\ \ell+c(u)>\lambda_i}} \frac{(d+1)m+\ell+c(u)}{\ell+c(u)}, \tag{7.194}$$

where $c(u)$ denotes the content of the square u.

7.102. **a.** [2–] For $\lambda \in \operatorname{Par}$, let n be large enough that $n + c(u) > 0$ for all $u \in \lambda$. (Specifically, $n \geq \ell(\lambda)$.) Define

$$t_{\lambda,n}(q) = s_\lambda(1, q, q^2, \ldots)\prod_{u\in\lambda}\left(1 - q^{n+c(u)}\right).$$

Show that $t_{\lambda,n}(q)$ is a polynomial in q with nonnegative integer coefficients.

b. [3−] Generalize (a) to skew shapes λ/μ. Here we define $c(u)$ for $u \in \lambda/\mu$ by restriction from λ. (For example, $21/1$ has contents 1 and -1, so we must take $n \geq 2$.) Thus if

$$t_{\lambda/\mu,n}(q) = s_{\lambda/\mu}(1, q, q^2, \ldots)\prod_{u\in\lambda/\mu}\left(1 - q^{n+c(u)}\right),$$

then show that $t_{\lambda/\mu,n}(q)$ is a polynomial in q with nonnegative integer coefficients. More precisely,

$$t_{\lambda/\mu,n} = \sum_T q^{|T|},$$

summed over all reverse SSYT $T = (T_{ij})$ (allowing 0 as a part) of shape λ/μ such that $T_{ij} \le n + \mu_i - i$. For instance, if $\lambda/\mu = 32/1$ and $n = 2$, then the tableaux enumerated by $t_{32/1,2}(q)$ are given by

*

$$\begin{matrix} 0\,0 \\ 0\,1 \end{matrix} \qquad \begin{matrix} 0\,0 \\ 1\,1 \end{matrix} \qquad \begin{matrix} 0\,0 \\ 0\,2 \end{matrix} \qquad \begin{matrix} 0\,0 \\ 1\,2 \end{matrix} \qquad \begin{matrix} 0\,0 \\ 2\,2 \end{matrix}.$$

Hence $t_{32/1,2}(q) = q + 2q^2 + q^3 + q^4$.

7.103. a. [3+] Let $A(r)$ be the number of plane partitions π with at most r rows such that π is symmetric and every row of π is a self-conjugate partition. (It follows that π has at most r columns and largest part at most r.) Such plane partitions are called *totally symmetric*. Show that

$$A(r) = \prod_{1 \le i \le j \le k \le r} \frac{i+j+k-1}{i+j+k-2}.$$

b. [3+] Let $B(r)$ be the number of plane partitions as in (a) which are also self-complementary (as defined in Exercise 7.106(b)). Show that

$$B(r) = \frac{1!\,4!\,7!\,10! \cdots (3r-2)!}{r!\,(r+1)!\,(r+2)! \cdots (2r-1)!}.$$

c. [4−] A *monotone triangle* of order r is a Gelfand–Tsetlin pattern (as defined in Section 7.10) with first row $1, 2, \ldots, r$, for which every row is *strictly* increasing. Let $M(r)$ be the number of monotone triangles of order r. For instance, $M(3) = 7$, corresponding to

$$\begin{matrix} 1\,2\,3 \\ 1\,2 \\ 1 \end{matrix} \quad \begin{matrix} 1\,2\,3 \\ 1\,2 \\ 2 \end{matrix} \quad \begin{matrix} 1\,2\,3 \\ 1\,3 \\ 1 \end{matrix} \quad \begin{matrix} 1\,2\,3 \\ 1\,3 \\ 2 \end{matrix} \quad \begin{matrix} 1\,2\,3 \\ 1\,3 \\ 3 \end{matrix} \quad \begin{matrix} 1\,2\,3 \\ 2\,3 \\ 2 \end{matrix} \quad \begin{matrix} 1\,2\,3 \\ 2\,3 \\ 3 \end{matrix}.$$

Show that $M(r) = B(r)$.

d. [3] Let P be a poset with $\hat{1}$. The *MacNeille completion* $L(P)$ of P (mentioned in the solution to Exercise 3.12) is the meet semilattice of 2^P (the boolean algebra of all subsets of P) that is generated by the principal order ideals of P. Let P_n denote the Bruhat order of the symmetric group $\mathfrak{S}_n$, as defined in Exercise 3.75(a). Show that $\#L(P_n) = M(n)$. Figure 7-19 shows $L(P_4)$, with the elements of P_4 indicated by open circles.

7.104. [3+] Write $f(n) \sim g(n)$ if $\lim_{n\to\infty} f(n)/g(n) = 1$. Let $a(n)$ denote the number of plane partitions of n. Show that

$$a(n) \sim \zeta(3)^{1/36} 2^{-11/36} n^{-25/36} \exp(3 \cdot 2^{-2/3} \zeta(3)^{1/3} n^{2/3} + 2C),$$

where ζ denotes the Riemann zeta function and

$$C = \int_0^\infty \frac{y \log y \, dy}{e^{2\pi y} - 1} = -0.0827105718 \cdots.$$

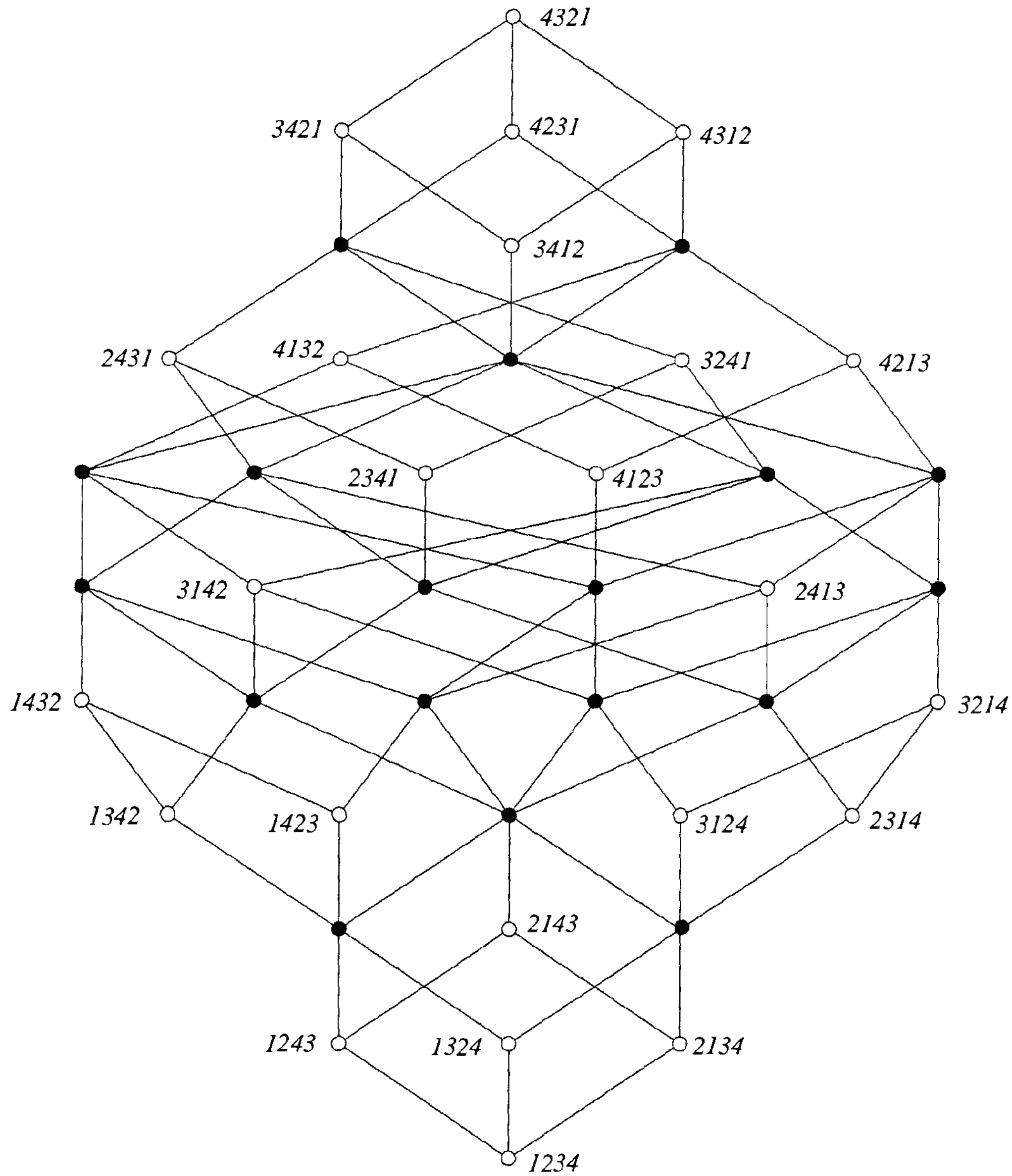


Figure 7-19. The MacNeille completion of the Bruhat order of $\mathfrak{S}_4$.

7.105. [3−] If the partitions λ and μ have the same multiset of hook lengths, does it follow that λ and μ are equal or conjugate?

7.106. a. [2] Let $\nu = \langle c^r \rangle$, the partition with r parts equal to c. Find the expansion of s_ν^2 in terms of Schur functions.

b. [3–] Fix r, c, and t. Let $\pi = (\pi_{ij})$ be a plane partition with at most r rows, at most c columns, and with largest part at most t. We say that π is *self-complementary*, or more precisely (r, c, t)-*self-complementary*, if $\pi_{ij} = t - \pi_{r-i,c-j}$ for all $1 \leq i \leq r$ and $1 \leq j \leq c$. In other words, π is invariant under replacing each entry k by $t - k$ (where we regard π as being an $r \times c$ rectangular array) and rotating 180°. For example, the following

where $a(k)$ is the the number of plane partitions of k, and $b_\mu(n-k)$ is the number of reverse plane partitions of $n-k$ of shape μ. Find a bijective proof.

7.108. [2−] Let $p, q \geq 2$. Find explicitly the number $F(p, q)$ of $w \in \mathfrak{S}_{p+q}$ with longest increasing subsequence of length p and longest decreasing subsequence of length q.

7.109. [3] Let $E(n)$ denote the expected length of the longest increasing subsequence of $w \in \mathfrak{S}_n$. Equivalently,

$$E(n) = \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} \mathrm{is}(w).$$

a. [2−] Show that

$$E(n) = \sum_{\lambda \vdash n} \lambda_1 (f^\lambda)^2. \tag{7.196}$$

b. [3] Show that $\lim_{n\to\infty} E(n)/\sqrt{n}$ exists.

c. [2] Let α denote the limit in (b). Assuming that α does indeed exist, deduce from Example 7.23.19(a) that $\alpha \geq 1$.

d. [3−] Show that $\alpha \leq e$.

e. [3+] Write $\tilde{\lambda}^n = ((\tilde{\lambda}^n)_1, (\tilde{\lambda}^n)_2, \ldots)$ for some partition of n that maximizes f^λ (over all $\lambda \vdash n$). Identify $\tilde{\lambda}^n$ with the function from $\mathbb{R}_{>0}$ to $\mathbb{R}_{\geq 0}$ defined by

$$\tilde{\lambda}^n(x) = \frac{(\tilde{\lambda}^n)_i}{\sqrt{n}} \qquad \text{if } \frac{i-1}{\sqrt{n}} < x \leq \frac{i}{\sqrt{n}}.$$

Thus $\int_0^\infty \tilde{\lambda}^n(x)\,dx = 1$. Show that for weak convergence in a certain "reasonable" metric, we have

$$\lim_{n\to\infty} \tilde{\lambda}^n = f,$$

where $y = f(x)$ is defined parametrically by

$$x = y + 2\cos\theta, \quad y = \frac{2}{\pi}(\sin\theta - \theta\cos\theta), \qquad 0 \leq \theta \leq 2\pi,$$

and $f(x) = 0$ for $x > 2$. Thus f describes the "limiting shape" of the partitions that maximize f^λ.

f. [3−] Deduce from (e) that $\alpha \geq 2$.

g. [3] Use the RSK algorithm to show that $\alpha \leq 2$. Hence $\alpha = 2$.

7.110. [3−] Let $d(T)$ denote the number of descents of the SYT T. Define

$$Z = \sum_{\lambda \in \mathrm{Par}} \left(\sum_{\substack{T \text{ is an SYT} \\ \text{of shape } \lambda}} q^{d(T)} \right) s_\lambda.$$

Show that

$$Z = \frac{\sum_{n\geq 0}(1-q)^n s_n}{1 - q\sum_{n\geq 1}(1-q)^{n-1} s_n}$$

plane partition is self-complementary for $(r, c, t) = (4, 5, 6)$:

$$\begin{bmatrix} 6 & 6 & 5 & 4 & 3 \\ 6 & 5 & 5 & 4 & 2 \\ 4 & 2 & 1 & 1 & 0 \\ 3 & 2 & 1 & 0 & 0 \end{bmatrix}$$

Let $F(r, c, t)$ denote the number of plane partitions with at most r rows, at most c columns, and with largest part at most t. Let $G(r, c, t)$ denote the number of such plane partitions that are self-complementary. (It is easy to see that $G(r, c, t) \neq 0$ if and only if rct is even.) Show that

$$G(2r, 2c, 2t) = F(r, c, t)^2. \tag{7.195}$$

Find similar formulas for $G(2r, 2c, 2t + 1)$ and $G(2r, 2c + 1, 2t + 1)$.

7.107. a. [2+] Let $\mu \in \mathrm{Par}$, and let A_μ be the infinite shape consisting of the quadrant $Q = \{(i, j) : i < 0,\ j > 0\}$ with the shape μ removed from the lower right-hand corner. Thus every square of A_μ has a finite hook and hence a hook length. For instance, when $\mu = (3, 1)$ we get the diagram

$$\begin{array}{c|c|c|c|c|c|c|}
 & & & & \vdots & & & \\
\hline
 & 10 & 9 & 8 & 6 & 5 & 3 \\
\hline
 & 9 & 8 & 7 & 5 & 4 & 2 \\
\hline
\cdots & 8 & 7 & 6 & 4 & 3 & 1 \\
\hline
 & 6 & 5 & 4 & 2 & 1 \\
\cline{2-6}
 & 3 & 2 & 1 \\
\cline{2-4}
\end{array}$$

Show that the multiset of hook lengths of A_μ is equal to the union of the multiset of hook lengths of Q (explicitly given by $\{1^1, 2^2, 3^3, \ldots\}$) and the multiset of hook lengths of μ.

b. [2+] Fix a plane partition μ, and let $a_\mu(n)$ be the number of skew plane partitions of n whose shape is λ/μ for some λ. For instance, $a_2(2) = 6$, corresponding to

$$\begin{array}{c} \cdot\cdot 2 \end{array} \qquad \begin{array}{l} \ \cdot\cdot \\ 2 \end{array} \qquad \begin{array}{c} \cdot\cdot 1\,1 \end{array} \qquad \begin{array}{l} \ \cdot\cdot 1 \\ 1 \end{array} \qquad \begin{array}{l} \ \cdot\cdot \\ 1\,1 \end{array} \qquad \begin{array}{l} \ \cdot\cdot \\ 1 \\ 1 \end{array}$$

Show that

$$\sum_{n \geq 0} a_\mu(n) q^n = \left(\prod_{i \geq 1} (1 - q^i)^{-i} \right) \left(\prod_{u \in \mu} \left(1 - q^{h(u)}\right)^{-1} \right).$$

c. [3] It follows from (b), Corollary 7.20.3, and Theorem 7.22.1 that

$$a_\mu(n) = \sum_{k=0}^{n} a(k) b_\mu(n - k),$$

and

$$Z = \sum_{\lambda} z_\lambda^{-1} q^{-1} (1-q)^{n-\ell} A_\ell(q)\, p_\lambda,$$

where $n = |\lambda|$, $\ell = \ell(\lambda)$, and $A_\ell(q)$ denotes an Eulerian polynomial.

7.111. Let B be a subset (called a *board*) of $[n] \times [n]$. Let $X = X_B$ be the set of all permutations $w \in \mathfrak{S}_n$ satisfying $w(i) = j \Rightarrow (i, j) \in B$. Let $\tilde{Z}_X(x)$ denote the augmented cycle indicator of X, as defined in Definition 7.24.1.

a. [2−] Let $B = [n] \times [n]$. Show that $\tilde{Z}_X = n!h_n$.

b. [2+] Let $X = \{w \in \mathfrak{S}_n : w(n) \neq 1\}$. Express $\tilde{Z}_X$ in terms of the basis $\{h_\lambda\}$.

c. [3] Suppose that there are integers $a, b \geq 0$ such that $a + b \leq n$ and $(i, j) \in B$ whenever $i \leq n - a$ or $j > b$. Let $m = \min\{a, b\}$. Show that $\tilde{Z}_X$ is a nonnegative (integer) linear combination of the symmetric functions $h_j h_{n-j}$, $0 \leq j \leq m$. (Note that (b) corresponds to the case $a = b = 1$.)

d. [5] Let $B \subseteq [n] \times [n]$, and suppose that the set $\{(i, n+1-j) : (i, j) \in B\}$ is the diagram of a partition. Show that $\tilde{Z}_X$ is *h-positive*.

e. [3−] Let $w \in \mathfrak{S}_n$, and let B_w denote the $n \times n$ chessboard with w removed, i.e.,

$$B_w = \{(i, j) \in [n]^2 : w(i) \neq j\}.$$

Show that

$$\tilde{Z}_{B_w} = \sum_{\lambda \vdash n} (f^\lambda)^{-1} d_\lambda \chi^\lambda(w) s_\lambda,$$

where d_λ is defined in Exercise 7.63(a).

f. [3−] Let w be an n-cycle in (e). Show that

$$\tilde{Z}_{B_w} = \sum_{i=1}^{n} [nD_{i-1} + (-1)^i] s_{\langle i, 1^{n-i} \rangle},$$

where D_{i-1} denotes the number of derangements of $[i-1]$.

7.112. **a.** [2+] Define two sequences $a_1 a_2 \cdots a_n$ and $b_1 b_2 \cdots b_n$ to be *equivalent* if one is a cyclic shift (conjugate) of the other. A *necklace* is an equivalence class of sequences. Show that the number $N(n, k)$ of necklaces of length n whose terms ("beads") belong to a k-element alphabet is given by

$$N(k, n) = \frac{1}{n} \sum_{d|n} \phi(d) k^{n/d}, \tag{7.197}$$

where ϕ denotes Euler's totient function.

b. [2+] Find a formula for the number of necklaces using n red beads and n blue beads (and no other beads).

7.113. [2+] Let $g_i(p)$ be the number of nonisomorphic graphs (without loops or multiple edges) with p vertices and i edges. Use Exercise 7.75(c) to show that the sequence $g_0(p), g_1(p), \ldots, g_{\binom{p}{2}}(p)$ is symmetric and unimodal.

Solutions to Exercises

7.1. True!
HINT: Where is the period at the end of the sentence?

7.2. See T. Brylawski, *Discrete Math.* **6** (1973), 201–219, and C. Greene and D. J. Kleitman, *Europ. J. Combinatorics* **7** (1986), 1–10. Note that Exercise 3.55 is concerned with the Möbius function of $\mathrm{Par}(n)$. (The answer to (c) is $n = 7$.)

7.3. Let ω be a primitive cube root of unity. Then

$$\begin{aligned}\prod_{i\geq 1}\left(1+x_i+x_i^2\right) &= \prod_{i\geq 1}(1-\omega x_i)(1-\omega^2 x_i)\\ &= \left(\sum_{n\geq 0}(-1)^n\omega^n e_n\right)\left(\sum_{n\geq 0}(-1)^n\omega^{2n}e_n\right)\\ &= \sum_n e_n^2 + \sum_{m<n}[(-1)^{m+n}\omega^{m+2n}+(-1)^{m+n}\omega^{2m+n}]e_m e_n\\ &= \sum_n e_n^2 + \sum_{m<n} c_{mn}e_m e_n,\end{aligned}$$

where

$$c_{mn} = \begin{cases} 2 & \text{if } m-n\equiv 0 \pmod 6\\ 1 & \text{if } m-n\equiv 1 \pmod 6\\ -1 & \text{if } m-n\equiv 2 \pmod 6\\ -2 & \text{if } m-n\equiv 3 \pmod 6\\ -1 & \text{if } m-n\equiv 4 \pmod 6\\ 1 & \text{if } m-n\equiv 5 \pmod 6).\end{cases}$$

This result is due to I. M. Gessel.

7.4. One of many ways to prove this formula (known to Jacobi) is to take the formula $s_\lambda = a_{\lambda+\delta}/a_\delta$ (Theorem 7.15.1), put $\lambda = (r)$, and expand the determinant $a_{(r)+\delta}$ by its last column. For further aspects, see R. A. Gustafson and S. C. Milne, *Advances in Math.* **48** (1983), 177–188.

7.5. By setting $y_1 = t$ and $y_2 = y_3 = \cdots = 0$ in (7.20) (or by reasoning directly from (7.11)), we get

$$\sum_{n\geq 1} p_n\frac{t^n}{n} = \log\sum_{n\geq 0} h_n t^n. \tag{7.198}$$

Differentiate with respect to t and multiply by t to get

$$\sum_{n\geq 1} p_n t^n = \frac{\sum_{n\geq 0} n h_n t^n}{\sum_{n\geq 0} h_n t^n}.$$

This is equivalent to the stated formula.

7.6. (Sketch.) Let $C_1, \ldots, C_j$ be the cycles of w of some fixed length i (so $j = m_i$). Choose a permutation $\pi \in \mathfrak{S}_j$ in $m_i!$ ways. Choose an element $a_k \in C_k$,

$1 \le k \le j$, in i^{m_i} ways. Do this for all i. Then there is a unique permutation v commuting with w such that if b_k is the least element of C_k, then $v(b_k) = a_{\pi(k)}$, and all v commuting with w are obtained in this way.

7.7. *Answer.* A basis consists of $\{p_\lambda : \lambda \vdash n,\ \text{all parts } \lambda_i > 0 \text{ are odd}\}$.

Proof. Clearly each such $p_\lambda \in \Omega^n$. Conversely, assume that $f \in \Omega^n$ but $f \notin \operatorname{span}_{\mathbb{Q}}\{p_\lambda : \lambda \vdash n\}$. We can assume that $f = \sum_\lambda c_\lambda p_\lambda$, where λ ranges over all partitions of n with at least one even part. Let μ be a *least* element in dominance order for which $c_\mu \neq 0$. Then the coefficient of $x_1^{\mu_1} x_3^{\mu_2} x_4^{\mu_3} \cdots$ in $f(x_1, -x_1, x_3, x_4, \ldots)$ is nonzero [why?], a contradiction. □

It follows that $\dim \Omega^n$ is the number of partitions of n into odd parts. By a famous theorem of Euler (see for instance item 10 of the Twelvefold Way in Section 1.4, as well as G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, fourth edition, Oxford University Press, London, 1960 (Thm. 344), and [**1.1**, Cor. 1.2]), this is also the number of partitions of n into distinct parts. A natural basis for Ω^n (due to Schur) analogous to Schur functions and indexed by partitions of n into distinct parts appears in [**96**, Ch. III.8].

7.8. Let $f = \sum_{\lambda \vdash n} c_\lambda p_\lambda$. Then $\omega f = \sum_{\lambda \vdash n} c_\lambda \varepsilon_\lambda p_\lambda$, so

$$(\omega f)_k = \sum_{\lambda \vdash n} c_\lambda \varepsilon_\lambda p_\lambda\big(x_1^k, x_2^k, \ldots\big) = \sum_{\lambda \vdash n} c_\lambda \varepsilon_\lambda p_{k\lambda}.$$

On the other hand,

$$f_k = \sum_{\lambda \vdash n} c_\lambda p_\lambda\big(x_1^k, x_2^k, \ldots\big) = \sum_{\lambda \vdash n} c_\lambda p_{k\lambda},$$

so $\omega f_k = \sum_{\lambda \vdash n} c_\lambda \varepsilon_{k\lambda} p_{k\lambda}$. Now

$$\begin{aligned} \varepsilon_{k\lambda} &= (-1)^{|k\lambda| - \ell(k\lambda)} = (-1)^{k\cdot|\lambda| - \ell(\lambda)} \\ &= (-1)^{(k-1)n}(-1)^{n-\ell(\lambda)} = (-1)^{(k-1)n}\varepsilon_\lambda, \end{aligned}$$

and the proof follows.

7.9. (Sketch.) From $a_{\lambda\mu} = \langle f_\lambda, h_\mu \rangle$ we get $h_\mu = \sum_\lambda a_{\lambda\mu} e_\lambda$. From (7.13) or otherwise, one shows that

$$h_n = \sum_{\lambda \vdash n} \varepsilon_\lambda c_\lambda e_\lambda,$$

where c_λ is the number of distinct permutations of $(\lambda_1, \ldots, \lambda_\ell)$ (where $\ell = \ell(\lambda)$) and hence is just the multinomial coefficient

$$\binom{\ell}{m_1(\lambda), m_2(\lambda), \ldots} = \frac{\ell!}{m_1(\lambda)! m_2(\lambda)! \cdots}.$$

Since h_λ and e_λ are multiplicative, one can compute $a_{\lambda\mu}$ by expanding $h_{\mu_1} h_{\mu_2} \cdots$ in terms of the e_λ's.

Forgotten symmetric functions were first "remembered" by P. Doubilet, *Studies in Applied Math.* **51** (1972), 377–396. A different (less straightforward) proof from that given above appears in this reference.

7.10. We have

$$\begin{aligned}\log A_\lambda(x) &= \sum_\alpha \log(1-x^\alpha)^{-1}\\ &= \sum_\alpha \sum_{n\geq 1} \frac{x^{\alpha n}}{n}\\ &= \sum_{n\geq 1} \frac{1}{n} m_\lambda\big(x_1^n, x_2^n, \dots\big)\\ &= \sum_{n\geq 1} \frac{1}{n} m_{n\lambda}.\end{aligned}$$

Suppose that $\lambda \vdash r$. Then [why?]

$$\omega m_\lambda\big(x_1^n, x_2^n, \dots\big) = (-1)^{(n-1)r}\varepsilon_\lambda \sum_\mu a_{\lambda\mu} m_\mu\big(x_1^n, x_2^n, \dots\big).$$

Hence

$$\begin{aligned}\omega \log A_\lambda(x) &= \log \omega A_\lambda(x)\\ &= \sum_{n\geq 1} \frac{1}{n}(-1)^{(n-1)r}\varepsilon_\lambda \sum_\mu a_{\lambda\mu} m_{n\mu}\\ &= (-1)^r \varepsilon_\lambda \sum_\mu \sum_{\beta\in \mathrm{Perm}(\mu)} \sum_{n\geq 1} \frac{(-1)^{nr}}{n} x^{\beta n},\end{aligned}$$

where $\mathrm{Perm}(\mu)$ denotes the set of distinct permutations of $(\mu_1, \mu_2, \dots)$. We get

$$\log \omega A_\lambda(x) = (-1)^r \varepsilon_\lambda \sum_\mu a_{\lambda\mu} \sum_\beta \log[1-(-1)^r x^\beta]^{-1},$$

so

$$\omega A_\lambda(x) = \begin{cases} \prod_\mu A_\mu(x)^{\varepsilon_\lambda a_{\lambda\mu}}, & r \text{ even}\\ \prod_\mu B_\mu(x)^{\varepsilon_\lambda a_{\lambda\mu}}, & r \text{ odd}.\end{cases}$$

Similarly (or because the transition matrix $M(f, m)$ is an involution),

$$\omega B_\lambda(x) = \begin{cases} \prod_\mu B_\mu(x)^{\varepsilon_\lambda a_{\lambda\mu}}, & r \text{ even}\\ \prod_\mu A_\mu(x)^{\varepsilon_\lambda a_{\lambda\mu}}, & r \text{ odd}.\end{cases}$$

7.11. *Answer:* $\sum_{j=0}^{n-1}(q-1)^j s_{n-j,1^j}$. Once this answer is guessed, it can be verified as follows. We obtain an SSYT of shape $(n-j, 1^j)$ and type μ by choosing which parts of μ, excluding the part 1, go in the j squares in the first column below the first row. There are $\binom{j}{\ell(\mu)-1}$ such choices, so $K_{(n-j,1^j),\mu} = \binom{j}{\ell(\mu)-1}$. Hence the coefficient of m_μ in the claimed answer is given by

$$\sum_{j=0}^{n-1} \binom{j}{\ell(\mu)-1}(q-1)^j = q^{\ell(\mu)-1},$$

and the proof follows.

7.12. This result was conjectured by E. Snapper, *J. Algebra* **19** (1971), 520–535 (Conjecture 9.1), and proved independently by R. A. Liebler and M. R. Vitale, *J. Algebra* **25** (1973), 487–489, and T. Y. Lam, *J. Pure Appl. Algebra* **10** (1977), 81–94 (Thm. 1).

7.13. **a.** See A. D. Berenshtein (= Berenstein) and A. V. Zelevinskii (= Zelevinsky), *Funct. Analysis Appl.* **24** (1990), 259–269; Russian original, 1–13.

7.14. **a.** By Corollary 7.13.7, the number in question is the number $S_3(r)$ of 3×3 symmetric $\mathbb{N}$-matrices for which every row sum is r. The desired formula is now an easy consequence of the expression for $G_3(\lambda)$ following Proposition 4.6.21.

b. Now by Corollary 7.13.7, we are just counting the number $S_n(r)$ of $n \times n$ symmetric $\mathbb{N}$-matrices for which every row sum is r. Proposition 4.6.21 shows that $S_n(r)$ has the form $P_n(r) + (-1)^r Q_n(r)$. It is not difficult to find deg $P_n(r)$, e.g., by arguing as in the proof of Proposition 4.6.19 or by computing the maximum number of linearly independent $n \times n$ symmetric $\mathbb{N}$-matrices. The value of deg $Q_n(r)$ is mentioned in the Notes to Chapter 4 as a conjecture. This conjecture was proved by Rong Qing Jia, in *Formal Power Series and Algebraic Combinatorics, Proceedings of the Fifth Conference, Florence, Italy, June 21–25, 1993* (A. Barlotti, M. Delest, and R. Pinzani, eds.), Università di Firenze, pp. 292–300, using the theory of multivariate splines. For a related paper, see R. Q. Jia, *Trans. Amer. Math. Soc.* **340** (1993), 179–198.

7.15. See I. G. Macdonald, *Bull. London Math. Soc.* **3** (1971), 189–192. For the case $p = 2$, see also J. McKay, *J. Algebra* **20** (1972), 416–418.

7.16. **a.** This result was first stated explicitly by E. A. Bender and D. E. Knuth, *J. Combinatorial Theory (A)* **13** (1972), 40–54. An earlier Pfaffian expression for a generalization of B_k was given by B. Gordon and L. Houten, *J. Combinatorial Theory* **4** (1968), 81–99. Gordon, *J. Combinatorial Theory* **11** (1971), 157–168, simplified a special case, which was equivalent to a specialization of B_k, to a determinant. Bender and Knuth observed that Gordon's simplification applied to B_k itself. Further discussion appears in I. M. Gessel, *J. Combinatorial Theory (A)* **53** (1990), 257–285 (§6).

b.–c. By Pieri's rule (Thm. 7.15.7), we get

$$s_{\lfloor n/2 \rfloor} s_{\lceil n/2 \rceil} = \sum_{\substack{\lambda \vdash n \\ \ell(\lambda) \leq 2}} s_\lambda .$$

Now take the coefficient of $x_1 x_2 \cdots x_n$ on both sides, and the formula for $y_2(n)$ follows. This argument is due to A. Regev, *Advances in Math.* **41** (1981), 115–136. Regev gives a similar (though more complicated) argument for $k = 3$. (See Exercise 7.82(b).) Combinatorial proofs for $2 \leq k \leq 5$ are due to D. Gouyou-Beauchamps, *Europ. J. Combinatorics* **10** (1989), 69–82. Gessel, *loc. cit.*, Thm. 15, deduces the formulas for $y_k(n)$, $2 \leq k \leq 5$, from (a).

d.–e. See Gessel, *loc. cit.*, §7. In this reference Gessel gives a slightly more complicated formula than (7.166) for $u_3(n)$, but he subsequently found the simplification stated here.

f. The case $k = 2$ was done in Exercise 6.19(xx). Some formulas for $y_k(n)$ and $u_k(n)$ for large k are given by I. P. Goulden, *Canad. J. Math.* **42** (1990), 763–774. For some related work (using Corollary 7.23.12), see Exercise 6.56(c).

7.17. See R. Stanley, *J. Combinatorial Theory* **76** (1996), 169–172. For further information on $W_i(n)$, see Exercise 6.33(c).

7.18. Label the Black pawns $P_1, \ldots, P_5$ from the bottom up. When pawn P_i promotes to a rook, call that rook R_i. Black's 25 moves are shown in Figure 7-20 as the elements of a poset P. Black can play his moves in any order such that if $u < v$ in P, then move u must precede move v. Hence the number of solutions is the number $e(P)$ of linear extensions of P. This number is just $f^{(6,6,6,6)}$, the number of SYTs of shape (6, 6, 6, 6), and the hook-length formula (Corollary 7.21.6) yields the answer 140,229,804. This problem was composed by K. Väisänen and appears (Problem 7) in the booklet *Queue Problems* cited in the solution to Exercise 6.23.

7.19. Given $a, b \geq 0$, consider those λ of the form given by Figure 7-21, so $\sigma = \lambda/\mu$ for some μ. Let $p_{a,b,\sigma}(n)$ be the number of such μ satisfying $|\mu| = n$. Then [why?]

$$\sum_{n\geq 0} p_{a,b,\sigma}(n+t)q^n = \frac{q^{ab+ar+bs}}{[a]!\,[b]!},$$

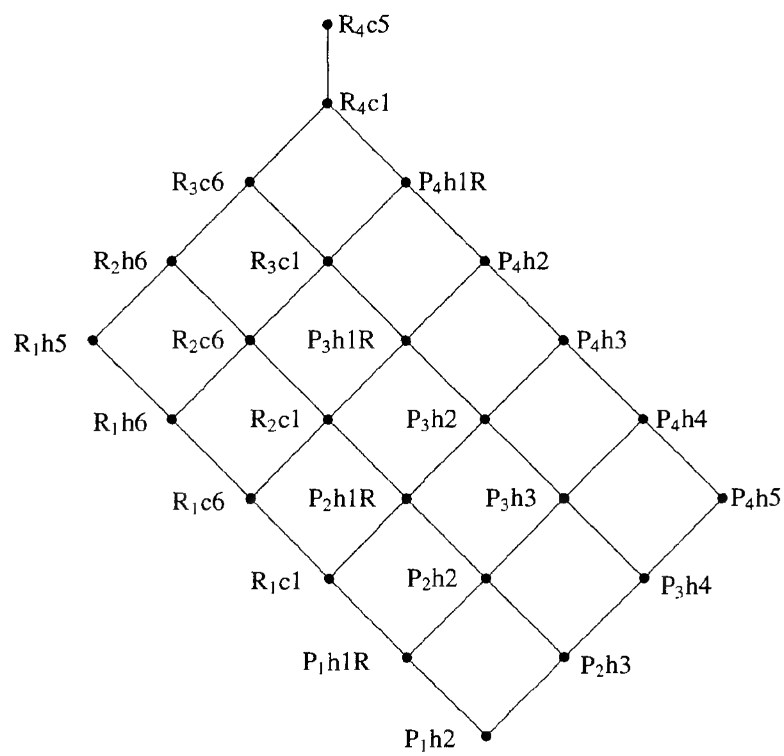


Figure 7-20. The solution poset to Exercise 7.18.

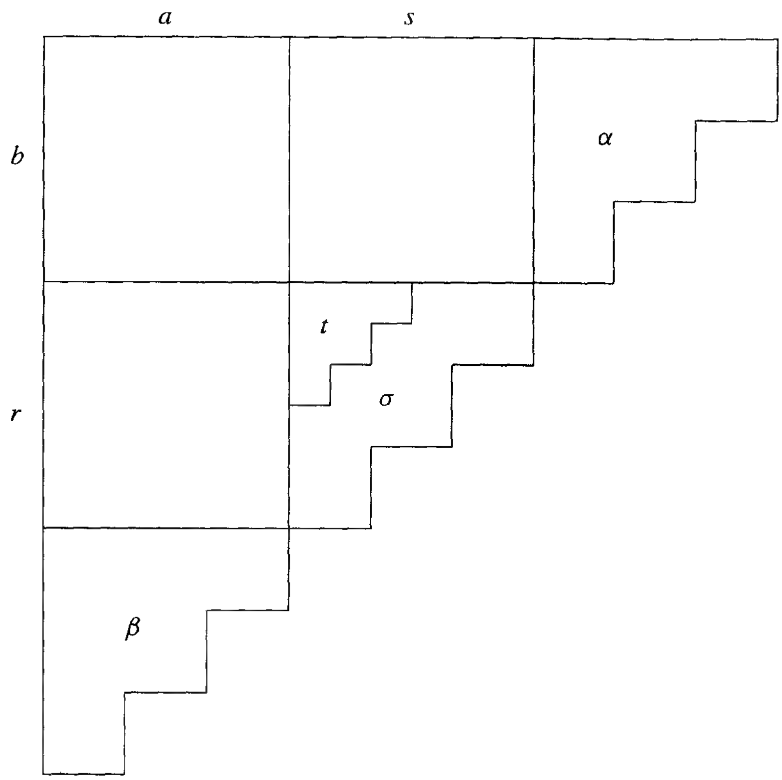


Figure 7-21. A skew shape λ/μ.

so

$$\sum_{n\geq 0} p_\sigma(n+t)q^n = \sum_{a,b\geq 0} \frac{q^{ab+ar+as}}{[a]!\,[b]!}.$$

Some simple manipulations show that the right-hand side is equal to the stated answer $[r-1]![s-1]!/[\infty]![r+s-1]!$.

7.20. **a.** In general, if $f \in \Lambda^n$ then $\langle p_1^n, f\rangle = [x_1\cdots x_n]f$, as follows e.g. from equation (7.25). Hence $\langle p_1^n, h_\lambda\rangle = [x_1\cdots x_n]h_\lambda = \binom{n}{\lambda_1,\lambda_2,\ldots}$. Since e.g. by Lemma 5.5.3 the number of partitions of $[n]$ of type λ is $\binom{n}{\lambda_1,\lambda_2,\ldots}\prod_i m_i(\lambda)^{-1}$, the result follows.

b. The following argument is due to Dale Worley. Let π be a partition of $[n]$ of type λ. Label the blocks $B_1, B_2, \ldots$ where $\#B_i = \lambda_i$, and if $\#B_i = \#B_j$ with $i < j$, then $\min B_i < \min B_j$. Insert the elements of B_i in increasing order into row i of λ. For instance, $B_1 = \{3,6,8\}$, $B_2 = \{5,7,9\}$, $B_3 = \{1,4\}$, $B_4 = \{2\}$ gives the array

$$\begin{matrix} 3\,6\,8 \\ 5\,7\,9 \\ 1\,4 \\ 2 \end{matrix}.$$

Sort each column into increasing order. For the above example, we get

$$\begin{array}{l} 1\,4\,8 \\ 2\,6\,9 \\ 3\,7 \\ 5 \end{array}.$$

The well-known "non-messing-up theorem" (see D. Gale and R. M. Karp, *J. Comput. System Sci.* **6** (1972), 103–115, for a more general result) states that the rows remain increasing, so an SYT T results. An easy combinatorial argument shows that the number of times a given SYT T occurs in this way is $f(T)$, and the proof follows.

c. Let T be an SSYT of shape λ and type μ. Let $T(s)$ denote the entry in the square $s \in \lambda$ of T. Call s *special* if $s = (i, j)$, $j > 1$, and $T(i, j-1) < T(i, j)$. If s is special, then define $f(s)$ exactly as in (b), i.e., $f(s)$ is the number of squares r in a column immediately to the left of s and in a row not above s, for which $T(r) < T(s)$. Now set

$$f(T) = \begin{cases} \prod_{\text{special}\, s} f(s) & \text{if } T \text{ has exactly } \ell(\mu) - \ell(\lambda) \text{ special squares} \\ 0 & \text{if } T \text{ has more than } \ell(\mu) - \ell(\lambda) \text{ special squares.} \end{cases}$$

(One can show that T always has at least $\ell(\mu) - \ell(\lambda)$ special squares.) Then

$$\sum_T f(T) = (m_i(\lambda)!)^{-1} \langle p_\mu, h_\lambda \rangle, \tag{7.199}$$

where T ranges over all SSYT of shape λ and type μ. The right-hand side of (7.199) is equal to the number of partitions of the multiset $\{1^{\mu_1}, 2^{\mu_2}, \ldots\}$ into *disjoint* blocks (where each block is a multiset) of sizes $\lambda_1, \lambda_2, \ldots$.

Example. Let $\lambda = (4, 2, 1)$ and $\mu = (2, 2, 1, 1, 1)$. There are five T with exactly $\ell(\lambda) - \ell(\mu) = 2$ special squares (whose entries are shown in boldface below), viz.,

$$\begin{array}{cccccc} T & \begin{array}{l} 11\mathbf{22} \\ 3\mathbf{4} \\ 5 \end{array} & \begin{array}{l} 11\mathbf{22} \\ 3\mathbf{5} \\ 4 \end{array} & \begin{array}{l} 11\mathbf{34} \\ 22 \\ 5 \end{array} & \begin{array}{l} 11\mathbf{35} \\ 22 \\ 4 \end{array} & \begin{array}{l} 11\mathbf{45} \\ 22 \\ 3 \end{array} \\ \\ f(T) & 1 & 2 & 2 & 2 & 2 \end{array}.$$

Thus $\sum_T f(T) = 9$, corresponding to the nine partitions of the multiset $\{1, 1, 2, 2, 3, 4, 5\}$ given by $1122-34-5$, $1122-35-4$, $1122-45-3$, $1134-22-5$, $1135-22-4$, $1145-22-3$, $2234-11-5$, $2235-11-4$, $2245-11-3$.

The proof of (c) is analogous to that of (b).

Note. Parts (b) and (c) were originally proved algebraically. Define $P_\lambda(x; t)$ as in [**96**, Ch. 3.2], and write

$$P_\lambda(x; t) = \sum_\mu (1-t)^{\ell(\lambda)-\ell(\mu)} \alpha_{\lambda\mu}(t) m_\mu(x),$$

where $\alpha_{\lambda\mu} \in \mathbb{Z}[t]$. One interprets $\alpha_{\lambda\mu}(1)$ in two ways, using (4.4) on p. 224 and (5.11′) on p. 229 of [**96**] (note that (4.4) has the typographical

error $Q_\lambda(x;t)$ instead of $Q_\lambda(y;t)$), and the proof follows. (The details are tedious.)

7.21. See P. H. Edelman and C. Greene, *Contemp. Math.* **34** (1984), 155–162 (Thm. 2), and *Advances in Math.* **63** (1987), 42–99 (Thm. 9.3). For a generalization of balanced tableaux, see S. Fomin, C. Greene, V. Reiner, and M. Shimozono, *Europ. J. Combinatorics* **18** (1997), 373–389.

7.22. a. For any $\alpha \in \mathrm{Comp}(p)$ we have $[x_1x_2\cdots x_p]L_\alpha = 1$, and the proof follows from the definition of F_w.

b. Define an algebra $\mathfrak{N}_n$ (over $\mathbb{Q}$, say), called the *nilCoxeter algebra* of the symmetric group $\mathfrak{S}_n$, as follows. $\mathfrak{N}_n$ has $n-1$ generators $u_1,\ldots,u_{n-1}$, subject to the *nilCoxeter relations*

$$\begin{aligned} u_i^2 &= 0, \qquad 1 \le i \le n \\ u_iu_j &= u_ju_i \qquad \text{if } |i-j| \ge 2 \\ u_iu_{i+1}u_i &= u_{i+1}u_iu_{i+1}, \qquad 1 \le i \le n-2. \end{aligned}$$

If $(a_1,\ldots,a_p) \in R(w)$, then identify the element $u_{a_1}\cdots u_{a_p}$ of $\mathfrak{N}_n$ with w. It is easy to see that this identification is well defined and that then $\mathfrak{S}_n$ is a $\mathbb{Q}$-basis for $\mathfrak{N}_n$. Write $\langle f, w\rangle$ for the coefficient of w when $f \in \mathfrak{N}_n$ is expanded in terms of the basis $\mathfrak{S}_n$.

Now let $\mathbf{x} = (x_1, x_2, \ldots)$ and define $A(x) \in \mathfrak{N}_n \otimes_{\mathbb{Q}} \mathbb{Q}[x]$ and $G = G(\mathbf{x}) \in \mathfrak{N}_n \otimes_{\mathbb{Q}} \mathbb{Q}[\mathbf{x}]$ by

$$\begin{aligned} A(x) &= (1+xu_{n-1})(1+xu_{n-2})\cdots(1+xu_1) \\ G &= A(x_1)A(x_2)\cdots. \end{aligned}$$

It is immediate from the definition of G that

$$G = \sum_{w\in\mathfrak{S}_n} F_{w^{-1}}\cdot w.$$

The crucial lemma, which has a simple proof by induction on n, asserts that

$$A(x)A(y) = A(y)A(x).$$

From this it follows that $F_w \in \Lambda^p$.

The result of this exercise was first given (with a more complicated proof) by R. Stanley, *Europ. J. Combinatorics* **5** (1984), 359–372 (Thm. 2.1). The proof sketched here appears in S. Fomin and R. Stanley, *Advances in Math.* **103** (1994), 196–207 (after Lemma 2.1).

c. Let $w = w_1w_2\cdots w_n$. Let a_w be the reduced decompostion of w obtained by starting with $12\cdots n$ and first moving w_n one step at a time to the last position, then w_{n-1} one step at a time to the next-to-last position, etc. For instance, if $w = 361524$ then $a_w = (4,5,2,3,4,3,1,2)$. One shows that $L_{\mathrm{co}(a_w)}$ contains the term $x_1^{\lambda_n(w)}x_2^{\lambda_{n-1}(w)}\cdots x_p^{\lambda_1(w)}$, that $L_{\mathrm{co}(a)}$ contains this term for no other $a \in R(w)$, and that no $L_{\mathrm{co}(a)}$ contains a term whose exponents, arranged in weakly decreasing order, are larger than $\lambda(w)$ in dominance order. Since $\sigma \le \rho$ whenever $K_{\sigma\rho} \ne 0$ and since $K_{\rho\rho} = 1$, we get that $\lambda \le \lambda(w)$ whenever $c_{w\lambda} \ne 0$ and that $c_{w,\lambda(w)} = 1$. The corresponding

results for $\mu(w)$ are a consequence of the results for $\lambda(w)$ applied to $F_{w^{-1}}$, together with the fact that $\omega F_w = F_{w^{-1}}$. (It is easy to deduce that $\omega F_w = F_{w^{-1}}$ from Exercise 7.94(a).) For details, see R. Stanley, *ibid.* (Thm. 4.1).

d. Vexillary permutations (though not yet with that name) were introduced by A. Lascoux and M. P. Schützenberger, *C. R. Acad. Sci. Paris, Série I* **294** (1982), 447–450 (see Thm. 3.1), and were independently discovered by R. Stanley, *ibid.* (Cor. 4.2). In the paper A. Lascoux and M. P. Schützenberger, *Letters in Math. Physics* **10** (1985), 111–124, vexillary permutations are *defined* to be 2143-avoiding permutations (p. 115), and the equivalence with the definition we have given is proved as Lemma 2.3. See also (1.27) of I. G. Macdonald, *Notes on Schubert Polynomials*, Publications du LACIM **6**, Université du Québec à Montréal, 1991.

e. This result is due to J. West, Ph.D. thesis, Massachusetts Institute of Technology, 1990 (Cor. 3.1.7), and *Discrete Math.* **146** (1995), 247–262 (Cor. 3.5). West gives a bijection between 2143-avoiding permutations and 4321-avoiding permutations in $\mathfrak{S}_n$. The proof then follows from the case $p = 3$ of Corollary 7.23.12 (replacing λ with λ').

f. The permutation w_0 is vexillary, and one easily sees that $\lambda_{w_0} = \mu_{w_0} = (n-1, n-2, \ldots, 1)$. Hence by (c) we get $r(w_0) = f^{n-1,n-2,\ldots,1}$, and the proof follows from the hook-length formula (Corollary 7.21.6). This result is due to R. Stanley, *ibid.* (Cor. 4.3.).

g. Formula (7.168) is a result of S. Fomin and A. N. Kirillov, *J. Algebraic Combinatorics* **6** (1997), 311–319 (Thm. 1.1). Notice that by (7.193) the product on the right-hand side of (7.168) is exactly the number of plane partitions of staircase shape $(n-1, n-2, \ldots, 1)$ with entries at most x. Formula (7.169) is due to Macdonald, *ibid.* (eqn. (6.11)). A simpler proof was given by S. Fomin and R. Stanley, *Advances in Math.* **103** (1994), 196–207 (Lemma 2.3). For a generalization, see Exercise 6.19(oo).

h. This result was first proved by P. H. Edelman and C. Greene, *Advances in Math.* **63** (1987), 42–99 (Cor. 8.4). For some subsequent proofs and related work, see W. Kraśkiewicz and P. Pragacz, Schubert functors and Schubert polynomials, preprint, October 1986, 22 pages; W. Kraśkiewicz and P. Pragacz, *C. R. Acad. Sci. Paris Ser. I Math.* **304** (1987), 209–211; W. Kraśkiewicz, *Europ. J. Combinatorics* **16** (1995), 293–313; S. Fomin and C. Greene, *Discrete Math.*, to appear (Thm. 1.2 and Example 2.2); V. Reiner and M. Shimozono, *J. Algebraic Combinatorics* **4** (1994), 331–351; and V. Reiner and M. Shimozono, *J. Combinatorial Theory (A)* **82** (1998), 1–73.

7.23. This surprising connection between the RSK algorithm and symmetric chain decompositions is due to K. P. Vo, *SIAM J. Algebraic Discrete Methods* 2 (1981), 324–332. For the special case when P is a boolean algebra, see also D. Stanton and D. White, *Constructive Combinatorics*, Springer-Verlag, New York, 1986 (Thm. 7.5).

7.24. a.–b. These are simple properties of differentiation having nothing to do with symmetric functions *per se*.

c. Straightforward proof by induction on ℓ.

d. As special cases of Theorem 7.15.7 and Corollary 7.15.9, together with the fact that $(\partial/\partial p_1)s_\lambda = s_{\lambda/1}$ (see the solution to Exercise 7.35(a)), we have

$$Us_\mu = \sum_\nu s_\nu, \qquad Ds_\mu = \sum_\rho s_\rho,$$

where ν is obtained from μ by adding a box, and ρ is obtained from μ by deleting a box. It follows easily that

$$(U+D)^\ell\, 1 = \sum_\lambda \tilde{f}^\lambda_\ell s_\lambda.$$

Since $D1 = 0$, it follows from (c) that

$$\begin{aligned}(U+D)^\ell\, 1 &= \sum_{\substack{i\leq\ell \\ r:=(n-i)/2\in\mathbb{N}}} \frac{\ell!}{2^r r!\, i!} U^i 1 \\ &= \sum_{\substack{i\leq\ell \\ r:=(n-i)/2\in\mathbb{N}}} \frac{\ell!}{2^r r!\, i!} \left(\sum_{\lambda\vdash i} f^\lambda s_\lambda\right),\end{aligned}$$

and the proof follows.

The operators U and D are powerful tools for enumerating various kinds of sequences obtained by adding and removing single squares of diagrams of partitions. Exercises 7.25–7.27 give some further examples. From the viewpoint of partially ordered sets, the fundamental property $DU - UD = I$ holds because Young's lattice Y is a *1-differential poset*, i.e., Y is a locally finite poset with $\hat{0}$ such that (i) if $\lambda \in Y$ covers exactly k elements, then λ is covered by exactly $k+1$ elements (see Exercise 3.22), and (ii) if distinct elements $\lambda, \mu \in Y$ cover exactly k common elements, then they are covered by exactly k common elements. (Note that in fact $k = 0$ or 1 in (ii).) The general theory of differential posets is developed in R. Stanley, *J. Amer. Math. Soc.* **1** (1988), 919–961 (see the top of p. 940 for the present exercise), and R. Stanley, in *Invariant Theory and Tableaux* (D. Stanton, ed.), IMA Vols. in Math. Appl. **19**, Springer-Verlag, New York, 1990, pp. 145–165. A generalization was given by I. M. Gessel, *J. Statist. Plann. Inference* **34** (1993), 125–134. Further references related to differential posets are S. Fomin, *J. Algebraic Combinatorics* **3** (1994), 357–404, and **4** (1995), 5–45; S. Fomin, *J. Combinatorial Theory (A)* **72** (1995), 277–292; R. Kemp, in *Proc. Fifth Conf. on Formal Power Series and Algebraic Combinatorics 1993*, pp. 71–80; D. Kremer and K. M. O'Hara, *J. Combinatorial Theory (A)* **78** (1997), 268–279; T. W. Roby, Ph.D. thesis, Massachusetts Institute of Technology, 1991; T. W. Roby, Schensted correspondences for differential posets, preprint; and R. Stanley, *Europ. J. Combin.* **11** (1990), 181–188.

e. We want [why?] a bijection between the set $\mathcal{O}^\ell_\lambda$ of oscillating tableaux of length ℓ ending at λ and pairs (π, T), where π is a partition of some subset S (necessarily of even cardinality) of $[\ell]$ into blocks of size two, and T is an SYT of shape λ on the letters $[\ell] - S$. Given an oscillating tableau $\emptyset = \lambda^0, \lambda^1, \ldots, \lambda^\ell = \lambda$, we will recursively define a sequence

$(\pi_0, T_0), (\pi_1, T_1), \ldots, (\pi_\ell, T_\ell)$ with $(\pi_\ell, T_\ell) = (\pi, T)$. We leave to the reader the task of verifying that this construction gives a correct bijection. Let π_0 be the empty partition (of the empty set $\emptyset$), and let T_0 be the empty SYT (on the empty alphabet). If $\lambda^i \supset \lambda^{i-1}$, then $\pi_i = \pi_{i-1}$ and T_i is obtained from T_{i-1} by adding the entry i in the square λ^i/λ^{i-1}. If $\lambda^i \subset \lambda^{i-1}$, then let T_i be the unique SYT (on a suitable alphabet) of shape λ^i such that T_{i-1} is obtained from T_i by column-inserting some number j. In this case let π_i be obtained from π_{i-1} by adding the block $B_i = \{i, j\}$.

This bijection is due to S. Sundaram, *J. Combinatorial Theory (A)* **53** (1990), 209–238. For connections between oscillating tableaux and representation theory, see S. Sundaram, in *Invariant Theory and Tableaux* (D. Stanton, ed.), IMA Vols. Math. Appl. **19**, Springer-Verlag, New York, 1990, pp. 191–225. For an approach to oscillating tableaux based on the growth diagrams of Section 7.13, see T. W. Roby, Schensted correspondences for differential posets, preprint (§4.2). For the theory of *skew* oscillating tableaux, see S. Dulucq and B. E. Sagan, *Discrete Math.* **139** (1995), 129–142, and T. W. Roby, *Discrete Math.* **139** (1995), 481–485. As an example of the above bijection (given by Sundaram in the first reference above), let the oscillating tableau be $(\emptyset,\ 1,\ 11,\ 21, 211, 111,\ 11,\ 21,\ 22,\ 221,\ 211)$. Then the pairs (B_i, T_i) (where B_i is the block added to π_{i-1} to obtain π_i) are given by

$$\begin{array}{cccccccccc}
\begin{matrix}1\end{matrix} & \begin{matrix}1\\2\end{matrix} & \begin{matrix}1\,3\\2\end{matrix} & \begin{matrix}1\,3\\2\\4\end{matrix} & \begin{matrix}1\\3\\4\end{matrix} & \begin{matrix}1\\3\end{matrix} & \begin{matrix}1\,7\\3\end{matrix} & \begin{matrix}1\,7\\3\end{matrix} & \begin{matrix}1\,7\\3\\9\end{matrix} & \begin{matrix}1\,7\\8\\9\end{matrix}\\[1em]
 & & & & \{2,5\} & \{4,6\} & & & & \{3,10\}.
\end{array}$$

Hence

$$T = \begin{matrix}1\,7\\8\\9\end{matrix}\ , \qquad \pi = \{\{2,5\},\{4,6\},\{3,10\}\}.$$

7.25. a. Let U and D be as in Exercise 7.24. It is easy to see that

$$f_{2k}(n) = \sum_{\lambda \vdash n} \langle (U+D)^{2k} s_\lambda, s_\lambda \rangle.$$

Now $U^i D^j(s_\lambda)$ is homogeneous of degree $n + i - j$. Hence setting

$$T = \sum_{i=0}^{k} \frac{(2k)!}{(k-i)!\, i!^2\, 2^{k-i}} U^i D^i,$$

we get by Exercise 7.24(c) that

$$\begin{aligned} f_{2k}(n) &= \sum_{\lambda \vdash n} \langle T s_\lambda, s_\lambda \rangle \\ &= \operatorname{tr}(T, \Lambda^n), \end{aligned}$$

where $\operatorname{tr}(T, \Lambda^n)$ denotes the trace of T acting on the space Λ^n. Note that

$$U^i D^i p_\mu = (m_1(\mu))_i\, p_\mu, \qquad (7.200$$

where $m_1(\mu)$ denotes the number of parts of μ equal to 1, and $(m_1(\mu))_i$ is the falling factorial. Since the trace of a linear transformation is the sum of its eigenvalues, we get

$$\begin{aligned} f_{2k}(n) &= \sum_{\mu \vdash n} \sum_{i=0}^{k} \frac{(2k)!}{(k-i)!\, i!^2\, 2^{k-i}} m_1(\mu)_i \\ &= \frac{(2k)!}{2^k\, k!} \sum_{\mu \vdash n} \sum_{i=0}^{k} \binom{m_1(\mu)}{i} \binom{k}{i} 2^i. \end{aligned}$$

Now, writing $P(q) = \prod_{j \geq 0}(1-q^j)^{-1}$, we have [why?]

$$\sum_{n \geq 0} \left[\sum_{\mu \vdash n} \binom{m_1(\mu)}{i} \right] q^n = \frac{q^i}{(1-q)^i} P(q).$$

Hence

$$\begin{aligned} \sum_{n \geq 0} f_{2k}(n) q^n &= \frac{(2k)!}{2^k k!} P(q) \sum_{i=0}^{k} \binom{k}{i} 2^i \frac{q^i}{(1-q)^i} \\ &= \frac{(2k)!}{2^k k!} P(q) \left(1 + \frac{2q}{1-q}\right)^k \\ &= \frac{(2k)!}{2^k k!} \left(\frac{1+q}{1-q}\right)^k P(q), \end{aligned}$$

completing the proof. This result appears (with a different proof, in the context of differential posets) in R. Stanley, *J. Amer. Math. Soc.* **1** (1988), 919–961 (Cor. 3.14).

b. It is easy to see that

$$\begin{aligned} g_{2k}(n) &= \sum_{\lambda \vdash n} \langle (UD)^k s_\lambda, s_\lambda \rangle \\ &= \mathrm{tr}((UD)^k, \Lambda^n). \end{aligned}$$

Hence if $\theta_1, \ldots, \theta_{p(n)}$ are the eigenvalues of UD acting on Λ^n, then

$$g_{2k}(n) = \theta_1^k + \cdots + \theta_{p(n)}^k.$$

It follows from the case $i = 1$ of (7.200) that the eigenvalues of UD are just the numbers $m_1(\mu)$, for $\mu \vdash n$. There are numerous ways to see that

$$\#\{\mu \vdash n : m_1(\mu) = n - j\} = p(j) - p(j-1),$$

and the proof follows. This result is related to R. Stanley, *J. Amer. Math. Soc.* **1** (1988), 919–961 (Thm. 4.1), and R. Stanley, in *Invariant Theory and Tableaux* (D. Stanton, ed.), IMA Vols. Math. Appl. **19**, Springer-Verlag, New York, 1990, pp. 145–165 (Prop. 2.9).

7.26. It is surprising that the only known proofs of this elementary identity involve either deep properties of Macdonald symmetric functions and q-Lagrange

inversion (A. M. Garsia and M. Haiman, *J. Algebraic Combinatorics* **5** (1996), 191–244 (Thm. 2.10(a))) or of the Hilbert scheme of points in the plane (M. Haiman, (t, q)-Catalan numbers and the Hilbert scheme, *Discrete Math.*, to appear (the case $m = 0$ of (1.10)). Naturally a more elementary proof would be desirable.

7.27. Algebraic proofs of (c)–(h) appear in [**96**, Exams. I.5.26–I.5.28] and were discovered independently by various persons (Lascoux, Towber, Stanley, Zelevinsky). The identities (a) and (b) are easily deduced from (f) and (h) by considering the exponential specialization of Section 7.8. All the identities (a)–(h) can also be proved using the operators U and D of Exercise 7.24. See R. Stanley, *J. Amer. Math. Soc.* **1** (1988), 919–961 (Thms. 3.2 and 3.11) for two cases. Finally, bijective proofs of (a)–(h), based on a skew generalization of the RSK algorithm, were given by B. E. Sagan and R. Stanley, *J. Combinatorial Theory (A)* **55** (1990), 161–193 (Cors. 4.5, 4.2, 6.12, 7.6, 6.4, 6.7, 7.4, and 6.9, respectively). Related work appears in S. Fomin, *J. Algebraic Combinatorics* **3** (1994), 357–404, and **4** (1995), 5–45; and *J. Combinatorial Theory (A)* **72** (1995), 277–292.

There is a special case of (c) that is especially interesting. Let $\beta = \emptyset$ and take the coefficient of $x_1 \cdots x_n$ on both sides. If $\alpha \vdash m$ then we obtain

$$\sum_{\lambda \vdash n} f^\lambda f^{\lambda/\alpha} = (n)_m f^\alpha.$$

In particular, if $m = n - 1$ then

$$\sum_{\lambda} f^\lambda = (m+1) f^\alpha, \tag{7.201}$$

where the sum on the left ranges over all partitions λ covering α in Young's lattice Y. This result has numerous other proofs, including a simple combinatorial argument using only the fact that if a partition μ covers k elements in Y, then it is covered by $k+1$ elements. Moreover, in terms of the character theory developed in Section 7.18, equation (7.201) asserts the obvious fact that

$$\dim \operatorname{ind}_{\mathfrak{S}_m}^{\mathfrak{S}_{m+1}} \chi^\alpha = (m+1) \dim \chi^\alpha.$$

7.28. **a.** As in the proof of Theorem 7.13.1, we may assume by Lemma 7.11.6 that if $\binom{u}{v}$ is the two-line array associated to A, then u and v have no repeated elements. The proof is by induction on the length n of u and v, the case $n = 0$ being trivial. Now (continuing the notation of the proof of Theorem 7.13.1) $\operatorname{tr}(A)$ is equal to the number of antichains $I_i\binom{u}{v}$ of odd cardinality [why?], and thus also equal to the number of antichains $I_i\binom{a}{b}$ of *even* cardinality. Thus by induction, $\operatorname{tr}(A)$ is the number of columns of $\bar{P}$ and $\bar{Q}$ of even length. Since the total number of columns of P and Q is the total number of antichains $I_i\binom{u}{v}$, the proof follows by induction. This result is due to M. P. Schützenberger [**140**, p. 127].

A proof can also be given based on the growth diagram of a permutation w used in the second proof of Theorem 7.13.1. Let $w \in \mathfrak{S}_n$ be an involution, so that the corresponding permutation matrix is symmetric. The

entire growth diagram $\mathcal{G}_w$ will then be symmetric. Let $\nu(i, j)$ be the partition appearing in square corner (i, j) of $\mathcal{G}_w$ (where the bottom left corner is $(0, 0)$). We claim that for $0 \leq i \leq n$, the number of columns of (the diagram of) $\nu(i, i)$ is equal to the number of fixed points k of w satisfying $k \leq i$. The proof is by induction on i, the case $i = 0$ being trivial. Assume the assertion for i. Let s_{ab} denote the square in the a-th row (from the bottom) and b-th column (from the left) of $\mathcal{G}_w$. By the induction hypothesis, we are assuming that the number of columns of $\nu(i, i)$ of odd length is equal to the number of fixed points k of w satisfying $k \leq i$. Consider the use of the local rules (L1)–(L4) to define $\nu(i+1, i+1)$. By the symmetry of $\mathcal{G}_w$ we have $\nu(i, i+1) = \nu(i+1, i)$, so rule (L3) never occurs. If (L1) applies, then $i+1$ is not a fixed point of w and $\nu(i+1, i+1) = \nu(i, i)$, as desired. If (L2) applies, then $i+1$ is not a fixed point of w, and $\nu(i+1, i+1)$ is obtained from $\nu(i, i)$ by adding 1 to two consecutive parts. This does not affect the number of columns of odd length, as desired. Finally, if (L4) applies then $i+1$ is a fixed point of w and $\nu(i+1, i+1)$ is obtained from $\nu(i, i)$ by adding a square to the first row. This increases by one the number of columns of odd length, as desired. Hence the proof follows by induction.

b. The left-hand side of (7.170) is equal to $\sum_A q^{\mathrm{tr}(A)} x^{\mathrm{row}(A)}$, where A ranges over all symmetric $\mathbb{N}$-matrices of finite support. Now use (a) together with Corollary 7.13.7.

c. Put $q = 0$ in (7.170). This identity was first proved by D. E. Littlewood [**88**, (11.9;2)] using symmetric functions. A bijective proof based on a version of the RSK algorithm was given by W. H. Burge [**13**, §2].

d. Since $f^\lambda = f^{\lambda'}$, we see that $a(n, k) = \sum_\lambda f^\lambda$ where λ' has k odd parts. By letting A be a symmetric permutation matrix in (a) (or by considering the coefficient of $q^k x_1 \cdots x_n$ in (7.170)) we get that $a(n, k)$ is the number of permutations in $\mathfrak{S}_n$ of cycle type $\langle 2^j, 1^k \rangle$, where $2j + k = n$. Hence

$$a(n, k) = \frac{n!}{2^j\, j!\, k!}.$$

e. *Answer:*

$$\prod_i (1 + q x_i)(1 - x_i^2)^{-1} \cdot \prod_{i<j} (1 - x_i x_j)^{-1} = \sum_\lambda q^{o(\lambda)} s_\lambda(x), \qquad (7.202)$$

where $o(\lambda)$ denotes the number of odd parts of λ.

The case $q = 0$ appears in [**88**, (11.9;4)]. A combinatorial proof was given by Burge, *ibid.* (§3). For noncombinatorial "modern" proofs of (a)–(e), see [**96**, Exams. I.5.4–I.5.10, pp. 76–79].

7.29. a. This result was first proved by D. E. Littlewood [**88**, (11.9;5) on p. 238]. A combinatorial proof was given by W. H. Burge, *ibid.* (§6). For a proof based on Weyl's denominator formula for the root system C_n, see [**96**, Exams. 9(c), pp. 78–79].

There is an interesting connection between the result of this exercise and algebraic topology. Define the *matching complex* M_n to be the simplicial complex whose vertices are the two-element subsets of $[n]$, and whose

faces consist of sets of pairwise disjoint vertices. The symmetric group $\mathfrak{S}_n$ acts naturally on M_n and hence also on its rational (reduced) homology $\tilde{H}_*(M_n;\ \mathbb{Q})$. We can therefore ask for the characteristic $\operatorname{ch}\tilde{H}_i(M_n;\mathbb{Q})$ of this action on the i-th homology group. Such a result was stated without proof (in a different form) by T. Józefiak and J. Weyman, *Math. Proc. Camb. Phil. Soc.* **103** (1988), 193–196 (p. 195). An explicit statement and proof was given by S. Bouc, *J. Algebra* **150** (1992), 158–186 (Proposition 4), and later independently by D. B. Karagueuzian, Ph.D. thesis, Stanford University, 1994. Namely,

$$\operatorname{ch}\tilde{H}_i(M_n;\mathbb{Q}) = \sum_{\lambda} s_\lambda, \tag{7.203}$$

where λ ranges over all self-conjugate partitions of n satisfying $i = \lfloor \frac{n-1}{2} \rfloor - \lceil \frac{1}{2}\operatorname{rank}(\lambda) \rceil$. (In particular, the action of $\mathfrak{S}_n$ on the entire homology $\tilde{H}_*(M_n;\mathbb{Q})$ has the elegant characteristic $\sum_{\substack{\lambda\vdash n\\ \lambda=\lambda'}} s_\lambda$.) Now the Hopf trace formula (see S. Sundaram, *Contemp. Math.* **178** (1994), 277–309, for a discussion of this technique) shows that

$$\sum_i (-1)^i \operatorname{ch} C_i(M_n;\mathbb{Q}) = \sum_i (-1)^i \operatorname{ch}\tilde{H}_i(M_n;\mathbb{Q}),$$

where $C_i(M_n;\mathbb{Q})$ denotes the space of (oriented) rational i-chains of M_n. The left-hand (respectively, right-hand) side corresponds to the degree n part of the left-hand (respectively, right-hand) side of equation (7.171). Thus (7.171) is equivalent to the computation of the $\mathfrak{S}_n$-equivariant Euler characteristic of M_n, while (7.203) is a refinement that gives the actual homology.

For further information on matching complexes and related complexes (including *chessboard complexes*, which are the analogues of M_n for complete bipartite graphs), see A. Björner, L. Lovász, S. T. Vrećica, and R. Živaljević, *J. London Math. Soc.* **49** (1994), 25–39 (§4); P. F. Garst, Ph.D. thesis, University of Wisconsin–Madison, 1979, 130 pp.; V. Reiner and J. Roberts, Minimal resolutions and the homology of matching and chessboard complexes, preprint, July 1997; and G. M. Ziegler, *Israel J. Math.* **87** (1994), 97–110. The work of Józefiak and Weyman, of Bouc, and of Karagueuzian discussed above computes the homology of matching complexes over $\mathbb{Q}$ (with the additional structure of an $\mathfrak{S}_n$-action). It is also interesting to consider their homology over $\mathbb{Z}$. Computations of Bouc, *ibid.*
* (§3.3) and E. Babson, A. Björner, S. Linusson, J. Shareshian, and V. Welker, Complexes of not i-connected graphs, MSRI Preprint No. 1997-054 (Table 3), suggest that torsion only occurs for the prime 3, but this question remains open.

b. This was proved using symmetric functions by D. E. Littlewood [**88**, (11.9;3)], and by a variation of the RSK algorithm by W. H. Burge, *ibid.* (§5).

c. This result is due to T. Józefiak and J. Weyman, *Advances in Math.* **56** (1985), 1–8. For another proof and a number of related results, see A. Lascoux and P. Pragacz, *J. Phys. A* **21** (1988), 4105–4114. A further reference is J. B. Remmel and M. Yang, *SIAM J. Discrete Math.* **4** (1991), 253–274.

7.30. **a.** We have $\lambda + \delta = d(\mu + \delta)$. Hence by Theorem 7.15.1 we get

$$\begin{aligned} s_\lambda(x_1, \ldots, x_n) &= \frac{a_{\lambda+\delta}(x_1, \ldots, x_n)}{a_\delta(x_1, \ldots x_n)} \\ &= \frac{a_{\mu+\delta}(x_1^d, \ldots x_n^d)}{a_\delta(x_1^d, \ldots, x_n^d)} \cdot \frac{a_\delta(x_1^d, \ldots, x_n^d)}{a_\delta(x_1, \ldots, x_n)} \\ &= s_\mu(x_1^d, \ldots, x_n^d) \prod_{i<j} \frac{x_i^d - x_j^d}{x_i - x_j}. \end{aligned}$$

b. Put $\mu = \emptyset$ in (a).

c. See A.-A. A. Jucis (=Yutsis), *Mat. Zametki* **27** (1980), 353–359, 492; and T. S. Sundquist, Ph.D. thesis, University of Minnesota, 1992 (pp. 49–52).

7.31. Let $0 \le a_1 < a_2 < \cdots < a_n \le p - 1$ and $0 \le b_1 < b_2 < \cdots < b_n \le p - 1$. These two sequences define the submatrix $B = [\zeta^{a_j b_k}]_{j,k=1}^n$. Let $x = (x_1, \ldots, x_n)$ and define the matrix $B(x) = [x_j^{b_k}]_{j,k=1}^n$, so $B = B(\zeta^{a_1}, \ldots, \zeta^{a_n})$. Let $\lambda = (b_n - n+1, b_{n-1} - n+2, \ldots, b_1)$. By Theorem 7.15.1 we have

$$\det B(x) = \pm s_\lambda(x) \prod_{1 \le j < k \le n} (x_j - x_k).$$

Since $\prod_{1 \le j < k \le n} (\zeta^{a_j} - \zeta^{a_k}) \ne 0$, we need to show that $s_\lambda(\zeta^{a_1}, \ldots, \zeta^{a_n}) \ne 0$. Suppose the contrary. Then $q = \zeta$ is a zero of the integer polynomial $s_\lambda(q^{a_1}, \ldots, q^{a_n})$, so

$$s_\lambda(q^{a_1}, \ldots, q^{a_n}) = L(q)(1 + q + \cdots + q^{p-1})$$

for some $L(q) \in \mathbb{Z}[q]$. Putting $q = 1$ gives $s_\lambda(1^n) \equiv 0 \pmod p$. But by equation (7.105) we have

$$s_\lambda(1^n) = \prod_{1 \le j < k \le n} \frac{b_k - b_j}{k - j} \not\equiv 0 \pmod p,$$

a contradiction.

This result was first proved (in a different way) by N. G. Čebotarev in 1948. For further proofs and references, see M. Newman, *Lin. Multilin. Algebra* **3** (1975/76), 259–262. The proof given here was found by R. Stanley in 1990.

7.32. **a.** Write the Schur functions appearing in (7.172) as quotients of determinants using Theorem 7.15.1. The numerators are transposes of each other, while the denominators can be evaluated from equation (7.105). This result, as well as the two examples in (b), is due to J. R. Stembridge (private communication). Stembridge's work was done in the more general context of characters of arbitrary complex semisimple Lie algebras.

b. Let $\mu = (1)$ and $\mu = \langle 21^{n-2} \rangle$, respectively, in (a).

7.33. **a.** It follows from Exercise 7.30(c) that $t(n)$ is the number of outdegree sequences $(\alpha_1, \ldots, \alpha_n)$ of a tournament on the vertex set $[n]$. Now use Exercise 4.32.

b. Let $t_k(n)$ be the number of distinct monomials appearing in $s_{k\delta}(x_1, \ldots, x_n)$. By a straightforward generalization of Exercise 7.30(c), $t_k(n)$ is the number of outdegree sequences of (loopless) directed graphs on n vertices with exactly k edges (ignoring direction) between any two distinct vertices. Applying Exercise 4.32(b) to the undirected graph Γ on n vertices with k edges between any two distinct vertices shows that $t_k(n)$ is the number of forests on n vertices whose edges are k-colored. Hence

$$\sum_{n\geq 0} t_k(n)\frac{x^n}{n!} = \exp \sum_{j\geq 1} j^{j-2}k^{j-1}\frac{x^j}{j!}.$$

7.34. See [**96**, Exams. 3.8, pp. 46–47]. This result is essentially due to Jacobi [**64**].

7.35. **a.** The easy way is to show that $D = \partial/\partial p_1$ (acting on polynomials in $p_1, p_2, \ldots$). See [**96**, Exam. I.5.3c, pp. 75–76]. One only needs to check that

$$\left\langle \frac{\partial}{\partial p_1} p_\lambda, p_\mu \right\rangle = \langle p_\lambda, p_1 p_\mu \rangle,$$

which is routine.

One can also give a direct combinatorial proof, based on the Littlewood–Richardson rule (Appendix 1, Section A1.3).

b. If D is *any* derivation on a ring R, then a simple formal computation shows that setting $[f, g] = (Df)g - f(Dg)$ defines a Lie algebra structure.

c. This identity first arose in the context of the Korteweg–deVries equation in M. Adler and J. Moser, *Commun. Math. Phys.* **61** (1978), 1–30. A combinatorial proof was given by I. P. Goulden, *Europ. J. Combinatorics* **9** (1988), 161–168. For additional information and references, see B. Leclerc, *Discrete Math.* **153** (1996), 213–227.

7.36. Use the fact that D_μ is adjoint to the homomorphism M_μ which multiplies by s_μ (Theorem 7.15.3). Since M_μ and M_ν commute, so do D_μ and D_ν.

7.37. **a.** We have $a_\delta = \det V_n$, where V_n denotes the Vandermonde matrix $(x_i^{n-j})_1^n$. Then $V_n^t V_n = (p_{2n-i-j})_1^n$, with the (temporary) convention $p_0 = n$. Take the determinant of both sides to get the expansion $a_\delta^2 = \det(p_{2n-i-j})$. This result is due to C. W. Borchardt, *Crelle's J.* **30** (1845), 38–45, and *J. de Math.* **12** (1845), 50–67.

b. In Exercise 7.42 set $m = n$ and $y_j = -qx_j$ to obtain

$$x_1 \cdots x_n (1-q)^n \prod_{i\neq j}(x_i - qx_j) = \sum_{\lambda\subseteq\langle n^n\rangle} (-q)^{\lambda} s_\lambda(x) s_{\tilde{\lambda}'}(x),$$

where $x = (x_1, \ldots, x_n)$. Now

$$\begin{aligned}
\left\langle s_\lambda s_{\tilde{\lambda}'}, s_{\langle n^n\rangle}\right\rangle_n &= \left\langle s_\lambda, s_{\langle n^n\rangle/\tilde{\lambda}'}\right\rangle_n && \text{by (7.60)}\\
&= \langle s_\lambda, s_{\lambda'}\rangle_n && \text{by Exercise 7.56(a)}\\
&= \begin{cases} 1, & \lambda = \lambda' \subseteq \langle n^n\rangle \\ 0 & \text{otherwise.}\end{cases}
\end{aligned}$$

(See D. E. Littlewood, *J. London Math. Soc.* **28** (1953), 494–500 (Thm. III), for a different argument.) Hence the coefficient of $s_{\langle n^n\rangle}$ in $x_1 \cdots x_n$

$(1-q)^n \prod_{i\neq j}(x_i - qx_j)$ is equal to

$$\sum_{\substack{\lambda\in\langle n^n\rangle \\ \lambda=\lambda'}} (-q)^{|\lambda|} = (1-q)(1-q^3)\cdots(1-q^{2n-1}) \qquad \text{[why?]}.$$

Divide by $(1-q)^n$ and set $q = 1$ to get that the coefficient of $s_{\langle n^n\rangle}$ in $x_1\cdots x_n\, a_\delta^2$ is equal to $(-1)^{\binom{n}{2}} 1\cdot 3\cdots(2n-1)$. (The sign $(-1)^{\binom{n}{2}}$ arises because $a_\delta^2 = (-1)^{\binom{n}{2}}\prod_{i\neq j}(x_i - x_j)$.) This is equivalent [why?] to the statement that the coefficient of $s_{\langle (n-1)^n\rangle}$ in a_δ^2 is also $(-1)^{\binom{n}{2}} 1\cdot 3\cdots(2n-1)$.

c. This result follows from Theorem 11.4 or Example 11.6(a) of J. R. Stembridge, Ph.D. thesis, Massachusetts Institute of Technology, 1985.

d. This is the case $q = 1$ of J. R. Stembridge, *Trans. Amer. Math. Soc.* **299** (1987), 319–350 (Cor. 6.2). A formula for c_λ is unknown in general. A further reference is A. N. Kirillov, *Adv. Ser. Math. Phys.* **16** (1992), 545–579. For some of the sophisticated algebra related to the symmetric function a_δ^2, see P. Hanlon, *Advances in Math.* **84** (1990), 91–134, and B. Kostant, *Advances in Math.* **125** (1997), 275–350 (see especially §5).

e. Set $F(q) = \prod_{i\neq j}(x_i - qx_j)$. Let $\omega = e^{2\pi i/3}$. Since

$$(x_i - \omega x_j)(x_j - \omega x_i) = -\omega\left(x_i^2 + x_i x_j + x_j^2\right),$$

we have

$$\begin{aligned} F(\omega) &= (-\omega)^{\binom{n}{2}} \prod_{i<j}\left(x_i^2 + x_i x_j + x_j^2\right) \\ &= (-\omega)^{\binom{n}{2}} s_{2\delta}(x_1,\ldots,x_n), \end{aligned}$$

using Exercise 7.30(b). By considering the largest exponent in dominance order of the monomials appearing in the expansion of $F(q)$, we see that

$$\langle F(q), s_{2\delta}\rangle = (-q)^{\binom{n}{2}}.$$

It follows that in the Schur function expansion of $F(q)$, all coefficients except that of $s_{2\delta}$ vanish at $q = \omega$. Hence (since these coefficients are polynomials with integer coefficients) they are divisible by $q^2 + q + 1$. Now put $q = 1$. This result was discovered empirically by J. Stembridge (private communication dated 13 May 1998) and proved by R. Stanley. For somewhat related results see [**96**, Ex. I.3.17, pp. 50–51].

7.38. a. For N sufficiently large (viz., $N \geq |\lambda/\mu|$), let V be a complex N-dimensional vector space, and let $F^{\lambda/\mu}$ denote a GL(V)-module with character $s_{\lambda/\mu}(x_1,\ldots,x_N)$ (using the results and terminology of Appendix 2). For a (weak) composition $\alpha = (\alpha_1,\alpha_2,\ldots)$, let S^α denote the GL(V)-module with character h_α (so $S^\alpha = S^{\alpha_1}(V)\otimes S^{\alpha_2}(V)\otimes\cdots$, where $S^i(V)$ denotes the i-th symmetric power of V). For $0 \leq j \leq \binom{n}{2}$, define the GL($V$)-module

$$J^j = \coprod_{\substack{w\in\mathfrak{S}_n \\ \ell(w)=j}} S^{\lambda+\delta-w(\mu+\delta)},$$

where $\coprod$ denotes direct sum, and where δ is as in equation (7.69). The idea

of the proof is to define an exact sequence (of GL(V)-modules)

$$0 \to J^{\binom{n}{2}} \to \cdots \to J^1 \to J^0 \to F^{\lambda/\mu} \to 0. \qquad (7.204)$$

The existence of such an exact sequence solves the problem at hand by an obvious extension of Exercise 2.3(b) from vector spaces to GL(V)-modules. The existence of the exact sequence (7.204) for the case $\mu = \emptyset$ was stated by A. Lascoux, Thèse, Université Paris VII, 1977, but without defining the maps. Actually, Lascoux considers the dual situation corresponding to Corollary 7.70, but the Schur positivity of $t_{\lambda/\mu,k}$ and that of $\omega(t_{\lambda/\mu,k})$ are equivalent. A rigorous treatment of Lascoux's work (for both the stated and the dual case) was subsequently given by K. Akin, *J. Algebra* **117** (1988), 494–503; in *Contemporary Math.* **88** (1989), pp. 209–217; and *J. Algebra* **152** (1992), 417–426. An independent treatment, for general λ/μ, was given by A. V. Zelevinskii (= Zelevinsky), *Functional Anal. Appl.* **21** (1987), 152–154.

7.39. This is a result of G. Z. Giambelli, *Atti Torino* **38** (1903), 823–844. See also [**96**, Exam. I.3.9, p. 47].

7.40. This is a result of A. Lascoux and P. Pragacz, *Europ. J. Combinatorics* **9** (1988), 561–574. See also [**96**, Exams. I.5.20–I.5.22, pp. 87–89].

7.41. *Algebraic proof.* By the classical definition of Schur functions (Theorem 7.15.1) we have

$$\begin{aligned}(x_1\cdots x_m)^n s_\lambda\left(x_1^{-1},\ldots,x_m^{-1}\right) &= \frac{(x_1\cdots x_m)^n \det\left(x_i^{-(\lambda_j+m-j)}\right)_1^m}{\det\left(x_i^{-(m-j)}\right)_1^m}\\ &= \frac{\det\left(x_i^{m+n-1}x_i^{-(\lambda_j+m-j)}\right)}{\det\left(x_i^{m-1}x_i^{-(m-j)}\right)}\\ &= \frac{\det\left(x_i^{n-\lambda_j+j-1}\right)}{\det\left(x_i^{j-1}\right)}\\ &= s_{\tilde{\lambda}}(x_1,\ldots,x_m).\end{aligned}$$

Combinatorial proof. Given an SSYT T of shape λ, let $\mu^1,\ldots,\mu^n$ be the columns of T (left to right). Let $\tilde{\mu}^i$ be the column whose entries are the complement in $[m]$ of the entries of μ^i, arranged in increasing order. Let $\tilde{T}$ have columns $\tilde{\mu}^n, \tilde{\mu}^{n-1},\ldots,\tilde{\mu}^1$. The map $T \mapsto T'$ yields the desired bijection. As an example, let $m = 6$, $n = 7$, and

$$T = \begin{matrix} 1\,1\,2\,3\,4 \\ 3\,3\,4\,4\,6 \\ 4\,5\,5 \\ 6 \end{matrix}.$$

Then

$$T' = \begin{array}{l} 1\,1\,1\,1\,1\,2\,2 \\ 2\,2\,2\,2\,3\,4\,5 \\ 3\,3\,3\,5\,6\,6 \\ 4\,4\,5\,6 \\ 5\,5 \\ 6\,6 \end{array} .$$

7.42. In the identity

$$\prod_{i=1}^{m}\prod_{j=1}^{n}(1+x_iy_j) = \sum_{\lambda} s_\lambda(x)s_{\lambda'}(y),$$

(obtained by specializing Theorem 7.14.3), replace y_j by y_j^{-1}, multiply by $(y_1\cdots y_n)^m$, and use Exercise 7.41. See [**96**, Exam. I.4(5), p. 67].

7.43. Note that

$$\psi(p_n) = \omega_y p_n(x,y)\big|_{\substack{x_1=1,\ y_1=t \\ x_i=y_i=0 \text{ if } i>0}},$$

where $x=(x_1,x_2,\ldots)$, $y=(y_1,y_2,\ldots)$, and ω_y denotes the involution ω acting on the y variables only. Hence (using equation (7.66)),

$$\begin{aligned} \psi(s_\lambda) &= s_\lambda(x,y)\big|_{\substack{x_1=1,\ y_1=t \\ x_i=y_i=0 \text{ if } i>0}} \\ &= \sum_{\mu\subseteq\lambda}\left(s_\mu(x)\Big|_{\substack{x_1=1 \\ x_i=0 \text{ if } i>0}}\right)\cdot\left(\omega_y s_{\lambda/\mu}\Big|_{\substack{y_1=t \\ y_i=0 \text{ if } i>0}}\right) \\ &= \sum_{\mu\subseteq\lambda} s_\mu(1)s_{\lambda'/\mu'}(t). \end{aligned}$$

Now $s_\mu(1)=0$ unless μ consists of a single row, in which case $s_\mu(1)=1$. Similarly, $s_{\lambda'/\mu'}(t)=0$ unless λ/μ is a vertical strip, in which case $s_{\lambda'/\mu'}(t)=t^{|\lambda/\mu|}$. Thus if $s_\mu(1)s_{\lambda'/\mu'}(t)\neq 0$, then $\lambda=\langle n-k,1^k\rangle$ for some $0\le k\le n-1$, and either $\mu=(n-k)$ or $\mu=(n-k-1)$, in which case $s_\mu(1)s_{\lambda'/\mu'}(t)=t^{|\lambda/\mu|}$. From this the proof is immediate.

Suppose that $f=\sum_{\mu\vdash n} c_\mu s_\mu$. Applying ψ and dividing by $1+t$ yields

$$\frac{1}{1+t}\psi(f) = \sum_{k=0}^{n-1} c_{\langle n-k,1^k\rangle}t^k.$$

Hence this exercise can be a useful tool for evaluating hook coefficients of Schur function expansions. See Exercise 7.86(c) for an example.

7.44. This interesting specialization is due to F. Brenti, *Pacific J. Math.* **157** (1993), 1–28. Parts (a)–(d) appear as Theorem 4.1, Theorem 4.2, Proposition 4.5, and Proposition 4.8, respectively.

7.45. Let $s_\lambda = \sum_{\mu \vdash ab} K_{\lambda\mu} m_\mu$, so $K_{\lambda\mu}$ is a Kostka number. Thus

$$T_a(s_\lambda) = \sum_{\nu \vdash b} K_{\lambda, a\nu} m_\nu.$$

If follows that for each $\rho \vdash b$ we have

$$\begin{aligned}
\langle s_\rho, T_a(s_\lambda)\rangle &= \left\langle s_\rho, \sum_{\nu \vdash b} K_{\lambda, a\nu} m_\nu \right\rangle \\
&= \left\langle s_\rho, \sum_\nu \langle s_\lambda, h_{a\nu}\rangle\, m_\nu \right\rangle \\
&= \sum_\nu \langle s_\rho, m_\nu\rangle \cdot \langle s_\lambda, h_{a\nu}\rangle \\
&= \left\langle s_\lambda, \sum_\nu \langle s_\rho, m_\nu\rangle\, h_{a\nu} \right\rangle,
\end{aligned}$$

using the bilinearity of the scalar product together with Corollary 7.12.4 and the orthonormality of Schur functions (Corollary 7.12.2).

Consider the algebra endomorphism φ_a of the ring Λ defined by $\varphi_a(h_i) = h_{ai}$. If we apply φ_a to the Jacobi–Trudi identity defining s_ρ (Theorem 7.16.1), then we obtain the Jacobi–Trudi matrix for the skew Schur function of skew shape $[a\rho + (a-1)\delta]/(a-1)\delta$, where if $\ell(\rho) = \ell$ then $\delta = (\ell - 1, \ell - 2, \ldots, 1, 0)$. Hence

$$\varphi_a(s_\rho) = s_{(a\rho+(a-1)\delta)/(a-1)\delta}.$$

Thus

$$\begin{aligned}
\sum_\nu \langle s_\rho, m_\nu\rangle h_{a\nu} &= \varphi_a\left(\sum_\nu \langle s_\rho, m_\nu\rangle h_\nu\right) \\
&= \varphi_a(s_\rho) \\
&= s_{(a\rho+(a-1)\delta)/(a-1)\delta}.
\end{aligned}$$

It follows that

$$\langle T_a(s_\lambda), s_\rho\rangle = \left\langle s_\lambda s_{(a-1)\delta}, s_{a\rho+(a-1)\delta}\right\rangle,$$

a Littlewood–Richardson coefficient. Such coefficients are always nonnegative (see Corollary 7.18.6 and Appendix 1, Section A1.3), and the proof follows.

This result is due independently to R. Stanley, *Electron. J. Combinatorics* **3**(2), R6 (1996), 22 pp. (Thm. 2.4), and to P. Littelmann, as a simple consequence of his path model theory developed in *Ann. Math.* **142** (1995), 499–525. More explicit statements appear in papers by Littelmann: *J. Amer. Math. Soc.* **11** (1998), 551–567 (§2); The path model, the quantum Frobenius map and Standard Monomial Theory, in *Algebraic Groups and Their Representations* (R. Carter and J. Saxl, eds.), Kluwer, Dordrecht, to appear.

7.46. This result was conjectured by C. Reutenauer, *Advances in Math.* **110** (1995), 234–246 (Conjecture 2), and proved independently by W. F. Doran, *J. Combinatorial Theory (A)* **74** (1996), 342–344, and by T. Scharf and J.-Y. Thibon, unpublished. Doran defines a symmetric function

$$f(n,k) = \sum_{\substack{\lambda \vdash n \\ \min\{\lambda_i : \lambda_i > 0\} \geq k}} q_\lambda$$

and shows that

$$-f(n,k) = s_{(n-1,1)} + \sum_{i=2}^{k} f(i,i) f(n-i,i).$$

It follows by induction that for $k \geq 2$ the symmetric function $-f(n,k)$ is a nonnegative linear combination of Schur functions. Since $q_n = f(n,n)$, the proof is complete.

7.47. **a.** The symmetric function X_G was first defined by R. Stanley, *Advances in Math.* **111** (1995), 166–194. (The reference in parentheses preceding the solutions to (b)–(g) and (j)–(k) below refers to the preceding reference.) The fact that $X_G(1^n) = \chi_G(n)$ is stated as Proposition 2.2, and is immediate from the definitions.

b. (p. 170) This question has been checked to be true for trees with at most nine vertices by T. Chow. Note that all trees with d vertices have the same chromatic polynomial, viz., $n(n-1)^{d-1}$.

c. (Proposition 2.4) The coefficient of a monomial $x_1^{\lambda_1} x_2^{\lambda_2} \cdots$ in X_G is equal to the number of ways to choose a stable partition π of G of type $\lambda = \langle 1^{r_1} 2^{r_2} \cdots \rangle$, and then for each i to color some block of size λ_i with the color i. Once we choose π we have $r_1! r_2! \cdots$ ways to choose the coloring, and the proof follows.

d. (Theorem 2.6) The solution is analogous to the solution of Exercise 3.44. By a *coloring* of G, we mean *any* map $\kappa : V \to \mathbb{P}$. (Note that "coloring" in Exercise 3.44 is here called "proper coloring.") Given $\sigma \in L_G$, define $X_\sigma = \sum_\kappa x^\kappa$, summed over all colorings κ of G such that (i) if u and v are in the same block of σ then $\kappa(u) = \kappa(v)$, and (ii) if u and v are in different blocks and there is an edge with vertices u and v, then $\kappa(u) \neq \kappa(v)$. Given *any* $\kappa : V \to \mathbb{P}$, there is a unique $\sigma \in L_G$ such that κ indexes one of the terms appearing in the definition of X_σ. It follows that for any $\pi \in L_G$ we have

$$p_{\text{type}(\pi)} = \sum_{\sigma \geq \pi} X_\sigma.$$

By Möbius inversion (Proposition 3.72),

$$X_\pi = \sum_{\sigma \geq \pi} p_{\text{type}(\sigma)} \mu(\pi, \sigma).$$

But $X_{\hat{0}} = X_G$, and the proof follows.

e. (Corollary 2.7) Since for any $\lambda \vdash d$ we have $\varepsilon_\lambda = (-1)^{d-\ell(\lambda)}$ (see equation (7.19)), there follows $\varepsilon_{\text{type}(\pi)} = (-1)^{d-|\pi|}$. Now use Proposition 3.10.1, Proposition 7.7.5, and equation (7.174).

f. (Theorem 3.1 and equation (8)) Let P be a d-element poset. Fix an order-reversing bijection $\omega : P \to [d]$. Let

$$X_P = \sum_\kappa x^\kappa,$$

summed over all *strict* order-preserving maps $\kappa : P \to \mathbb{P}$. By Corollary 7.19.5, we have

$$X_P = \sum_{\pi \in \mathcal{L}(P,\omega)} L_{\mathrm{co}(\pi)}, \tag{7.205}$$

so in particular X_P is L-positive. Now let $\mathfrak{o}$ be an acyclic orientation of G and κ a proper coloring. We say that κ is $\mathfrak{o}$-*compatible* if $\kappa(u) < \kappa(v)$ whenever (v, u) is an edge of $\mathfrak{o}$ (i.e., the edge uv of G is directed from v to u). It is easy to see that every proper coloring κ is compatible with exactly one acyclic orientation $\mathfrak{o}$. Hence if $X_\mathfrak{o} = \sum_\kappa x^\kappa$, summed over all $\mathfrak{o}$-compatible proper colorings κ, then $X_G = \sum_\mathfrak{o} X_\mathfrak{o}$, summed over all acyclic orientations of G. Let $\bar{\mathfrak{o}}$ denote the transitive and reflexive closure of $\mathfrak{o}$. Since $\mathfrak{o}$ is acyclic, $\bar{\mathfrak{o}}$ is a poset and $X_{\bar{\mathfrak{o}}} = X_\mathfrak{o}$. Since $X_{\bar{\mathfrak{o}}}$ is L-positive by (7.205), it follows that X_G is L-positive.

g. (Theorem 3.3) The idea of the proof is to define (using the notation of Section 7.19) a linear operator $\varphi : \mathcal{Q}^d \to \mathbb{Q}[t]$ by

$$\varphi(L_\alpha) = \begin{cases} t(t-1)^i, & \alpha = (i+1, 1, 1, \ldots, 1) \\ 0 & \text{otherwise.} \end{cases}$$

One then shows that $\varphi(e_\lambda) = t^{\ell(\lambda)}$ and

$$\varphi(X_G) = \sum_j \mathrm{sink}(G, j)\, t^j,$$

from which the proof follows. It would be interesting to have a more conceptual proof.

h. This beautiful result is due to V. N. Gasharov, *Discrete Math.* **157** (1996), 193–197, using an involution principle argument. For the case when P is a chain, a bijective proof follows immediately from the RSK algorithm. Thus it is natural to ask for a generalization of the RSK algorithm that proves the general case of Gasharov's result. When the poset P also contains no induced subposet isomorphic to the poset of Figure 7-22., such a generalization was given by T. S. Sundquist, D. G. Wagner, and J. West, *J. Combinatorial Theory (A)* **79** (1997), 36–52.

i. Let V be the vertex set of $G = \mathrm{inc}(P)$. If $\alpha : V \to \mathbb{N}$, then define G^α to be the graph obtained from G by replacing each vertex v of G by a clique (complete subgraph) $K_{\alpha(v)}$ of size $\alpha(v)$, and placing edges connecting every vertex of $K_{\alpha(v)}$ to every vertex of $K_{\alpha(u)}$ if uv is an edge of G. (If $\alpha(v) = 0$ then we are simply deleting the vertex v.) It follows from the definition of X_{G^α} that

$$\prod_i C(x_i) = \sum_{\alpha : V \to \mathbb{N}} X_{G^\alpha}.$$

It is easy to see that each G^α is the incomparability graph of a $(\mathbf{3}+\mathbf{1})$-free

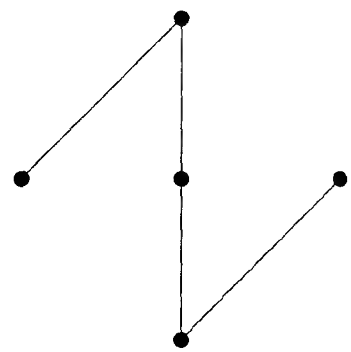


Figure 7-22. An obstruction to a generalized RSK algorithm.

poset. Hence by (h) the product $\prod_i C(x_i)$ is s-positive. The result now follows from Exercise 7.91(e). This argument appears in R. Stanley, Graph colorings and related symmetric functions: ideas and applications, *Discrete Math.*, to appear (Cor. 2.9). A different proof was given by M. Skandera in 1998.

j. (Conjecture 5.1) An equivalent conjecture appeared in R. Stanley and J. R. Stembridge, *J. Combinatorial Theory (A)* **62** (1993), 261–279 (Conjecture 5.5), in the context of "immanants of Jacobi–Trudi matrices." A special case of the conjecture asserts that for any fixed $k, d \geq 1$, the symmetric function $F_{k,d} = \sum x_{i_1} x_{i_2} \cdots x_{i_d}$, is e-positive, where the sum ranges over all sequences $i_1, i_2, \ldots, i_d$ such that any k consecutive terms are all distinct. Even the case $k = 3$ remains open. For the case $k = 2$, see equation (7.175).

k. (Propositions 5.3 and 5.4) Let $\lambda \vdash d$. The number b_λ of connected partitions of P_d (as defined in (d)) of type λ is just the number of distinct permutations of the parts of λ. Hence if $\lambda = \langle 1^{r_1} 2^{r_2} \cdots \rangle$, then $b_\lambda = \binom{\ell(\lambda)}{r_1, r_2, \ldots}$. Since P_d is a tree, the lattice L_{P_d} is just a boolean algebra, so by Example 3.8.3 we get $\mu(\hat{0}, \pi) = \varepsilon_{\text{type}(\pi)}$. Hence from equation (7.174) there follows

$$X_{P_d} = \sum_{\lambda \vdash d} \varepsilon_\lambda \binom{\ell(\lambda)}{r_1, r_2, \ldots} p_\lambda,$$

so

$$\sum_{d \geq 0} X_{P_d} \cdot t^d = \frac{1}{1 - p_1 t + p_2 t^2 - p_3 t^3 + \cdots}.$$

The proof now follows by applying the involution ω to equation (7.165).

A second proof appears in the reference given in (a). The result (stated in a different form) seems first to have been proved by L. Carlitz, R. A. Scoville, and T. Vaughan, *Manuscripta Math.* **19** (1976), 211–243 (p. 242). A combinatorial proof was given by J. Dollhopf, I. P. Goulden, and C. Greene, in preparation. The generating function (7.175) (or its image under the involution ω) appears in (seemingly) completely unrelated contexts in R. Stanley, in *Graph Theory and Its Applications: East and West*, Ann.

New York Acad. Sci. **576**, 1989, pp. 500–535 (based on work of C. Procesi; for further references see J. R. Stembridge, *Advances in Math.* **106** (1994), 244–301 (p. 266)), and in [**6.28**, Thm. 14.2.4].

The formula for C_d can be deduced from that of P_d by using Corollary 4.7.3. See p. 190 of the reference in (a) for details.

l. In fact, X_G is a nonnegative linear combination of the symmetric functions $e_i e_{d-i}$. In the special case when the complement of G is bipartite, this result follows from R. Stanley and J. R. Stembridge, *ibid.*, Theorem 7.4.3, and was given a different proof in R. Stanley, *Advances in Math.* **111** (1995), 166–194 (Cor. 3.6). It was observed by T. Chow that this second proof only requires that the complement of G be triangle-free.

* **m.** This question is due to V. N. Gasharov and appears in R. Stanley, Graph colorings and related symmetric functions: ideas and applications, *Discrete Math.*, to appear (Conjecture 1.4). If this question has an affirmative answer, then the following conjecture (first mentioned as an open question by Y. O. Hamidoune, *J. Combinatorial Theory (B)* **50** (1990), 241–244 (p. 242)) would follow in the same manner as (i). Let G be a clawfree graph, and let c_i be the number of stable i-element subsets of the vertex set of G. Then every zero of the polynomial $\sum c_i t^i$ is real.

7.48. The theory of locally rank-symmetric posets developed in this exercise first appeared in R. Stanley, *Electron. J. Combinatorics* **3**, R6 (1996), 22 pp. The definition of F_P was suggested by R. Ehrenborg, *Advances in Math.* **119** (1996), 1–25 (Def. 4.1).

a. By considering the support of the multichain $\hat{0} = t_0 \leq t_1 \leq \cdots \leq t_{k-1} < t_k = \hat{1}$, we get

$$F_P = \sum_{\substack{S=\{m_1,\ldots,m_j\}_< \\ S\subseteq[n-1]}} \sum_{1\leq i_1<\cdots<i_{j+1}} x_{i_1}^{m_1} x_{i_2}^{m_2-m_1} \cdots x_{i_{j+1}}^{n-m_j}\, \alpha_P(S), \tag{7.206}$$

where $\alpha_P(S)$ is defined in Section 3.12 (and is now called the *flag f-vector* of P). Since $\alpha_P(S) = \sum_{T\subseteq S} \beta_P(T)$ (equation (3.33)), we need to show that for each $T \subseteq [n-1]$,

$$\sum_{\substack{S\supseteq T \\ S=\{m_1,\ldots,m_j\}_<}} \sum_{1\leq i_1<\cdots<i_{j+1}} x_{i_1}^{m_1} x_{i_2}^{m_2-m_1} \cdots x_{i_{j+1}}^{n-m_j} = L_{\mathrm{co}(T)}.$$

But this is a routine verification, looking at all possible ways of choosing each symbol $\leq$ to be either $<$ or $=$ in the definition (7.89) of L_S. See R. Stanley, *ibid.* (Prop. 1.3).

b. This result is the special case $f = \zeta$ of Exercise 7.94(b). For a simplified version of the proof of Exercise 7.94(b) for the case at hand, see R. Simion and R. Stanley, *Discrete Math.* **204** (1999), 369–396 (Proposition 4.7.1).

c. It follows from (7.206) that $F_P \in \Lambda^n$ if and only if for all $S = \{m_1, \ldots, m_j\}_< \subseteq [n-1]$, we have that $\alpha_S(P)$ depends only on the multiset $\{m_1, m_2 - m_1, m_3 - m_2, \ldots, n - m_j\}$, not on the order of its elements R. Stanley, *Electron. J. Combinatorics*, **3**, R6 (1996), 22 pp., Cor. 1.2. Now use Exercise 3.65 (which applies to locally rank-symmetric posets, though it is stated only for locally self-dual posets).

d. *Ibid.* (Prop. 3.3).

e. *Ibid.* (Thm. 3.5). This result holds for a class of lattices, known as *q-semiprimary lattices of type μ*, more general than the subgroup lattices stated here. See the reference for further details. The polynomials $K_{\lambda\mu}(q) = q^{b(\lambda)}\tilde{K}_{\lambda\mu}(1/q)$ (where $b(\lambda) = \sum(i-1)\lambda_i$ as usual) are known as *Kostka polynomials* or *Kostka–Foulkes polynomials*. They occur in a variety of combinatorial and algebraic contexts and have many fascinating properties. For some further information on Kostka polynomials, see e.g. [**96**, Ch. 3.6], as well as A. M. Garsia and C. Procesi, *Advances in Math.* **94** (1992), 82–138; G.-N. Han, *Prépubl. Inst. Rech. Math. Av.* 1994/21, 71–85; and S. V. Kerov, A. N. Kirillov, and N. Yu. Reshetikhin, *Zap. Nauchn. Sem. Leningrad Otdel Mat. Inst. Steklov (LOMI)* **155** (1986), 50–64, 193.

f. See R. Stanley, *Electron. J. Combinatorics* **4**, R20 (1997), 14 pp. Some generalizations of NC_{n+1} related to root systems are discussed in §5 of that reference.

* **g.** See the reference cited in (b). Shuffle posets were first considered by C. Greene, *J. Combinatorial Theory (A)* **47** (1988), 191–206. A generalization was given by W. F. Doran, *J. Combinatorial Theory (A)* **66** (1994), 118–136.

7.49. This result was proved by C. Lenart, Lagrange inversion and Schur functions, preprint, 1998, by an intersecting lattice path argument.

7.50. In equation (7.78), set $x_1 = \cdots = x_n = 1$ and $x_{n+1} = x_{n+2} = \cdots = 0$ to get

$$s_\lambda(1^n) = \frac{1}{N!}\sum_{w\in\mathfrak{S}_N}\chi^\lambda(w)n^{c(w)}.$$

Now use Corollary 7.21.4 to get

$$\frac{1}{N!}\sum_{w\in\mathfrak{S}_N}\chi^\lambda(w)n^{c(w)} = \prod_{u\in\lambda}\frac{n+c(u)}{h(u)}. \tag{7.207}$$

(Of course a polynomial identity holding for all $n\in\mathbb{P}$ holds everywhere, i.e., when n is an indeterminate.)

7.51. Take the coefficient of n^{N-1} on both sides of equation (7.207), and the proof follows after some simple manipulation using the fact that $\sum_{u\in\lambda}c(u) = b(\lambda') - b(\lambda)$ (see [**96**, Exam. 3, p. 11]). This elegant proof is due to S. Sundaram and others. For other proofs, see [**96**, Exam. 7, pp. 117–118] and W. M. Benyon and G. Lusztig, *Math. Proc. Camb. Phil. Soc.* **84** (1978), 417–426 (pp. 419–420). One can also use Exercise 7.62 to get $\chi^\lambda(21^{n-2}) = f^{\lambda/2} - f^{\lambda/11}$. Clearly $f^{\lambda/2} + f^{\lambda/11} = f^\lambda$. To compute $f^{\lambda/2}$, see [**72**, Exer. 19, p. 70].

7.52. Define the *depth* $d(u)$ of a square $u = (i,j)$ of (the diagram of) λ to be the smallest integer $k > 0$ for which $u + (k,k) \notin \lambda$. The number of squares of depth k is just μ_k. Moreover, if we successively remove border strips from λ, then the k-th border strip removed contains no squares of depth greater than k. Hence the first k border strips removed contain a total of no more than $\mu_1 + \cdots + \mu_k$ squares. It is then immediate from the Murnaghan–Nakayama rule (Corollary 7.17.5) that if $\chi^\lambda(\nu) \neq 0$, then $\nu \leq \mu$. Now λ contains a unique border strip B_1 of size μ_1; $\lambda - B_1$ contains a unique border strip B_2 of size μ_2, etc. Since $\mathrm{ht}(B_i) = \lambda_i' - i$, we obtain from the Murnaghan–Nakayama rule the stated formula $\chi^\lambda(\mu) = (-1)^t$.

7.53. Consider equation (7.207). The coefficient of n^r on the left-hand side is equal to $1/N!$ times the sum $\sum_w \chi^\lambda(w)$ we are trying to evaluate. If now $\ell(\nu) > r$, then by Exercise 7.52 we have $\chi^\lambda(\nu) = 0$. Hence the left-hand side of (7.207) is divisible by n^r. There are r factors equal to n in the numerator of the right-hand side, coming from the r diagonal terms $u = (i, i)$. Hence the coefficient of n^r on the right-hand side is given by

$$\frac{\prod_{\substack{u\in\lambda\\ u\neq(i,i)}} c(u)}{\prod_{u\in\lambda} h(u)} = \frac{f^\lambda}{N!} \prod_{\substack{u\in\lambda\\ u\neq(i,i)}} c(u).$$

It is easy to see that this last product is just $(-1)^{t(\lambda)} \prod_{i=1}^r (\lambda_i - 1)!(\lambda_i' - 1)!$, and the proof follows.

7.54. Assume that $\chi^\lambda(\mu) = 0$ whenever some nonzero μ_i is even. Since $\chi^{\lambda'}(\mu) = \varepsilon_\lambda \chi^\lambda(\mu)$, it follows that $\lambda = \lambda'$. If $\lambda \neq (m, m-1, \ldots, 1)$, then λ has a border strip of even length. Let B be a maximum-size border strip of even length. Since any even border strip of a self-conjugate partition can be extended either up to the first row or down to the first column and remain even, it follows that there exist exactly two maximum-size even border strips B, B', symmetrically placed on the boundary of λ, and $\mathrm{ht}(B) \not\equiv \mathrm{ht}(B')\ (\mathrm{mod}\ 2)$. Let $|B| = 2r$. Then (by the Murnaghan–Nakayama rule) for all $\nu \vdash n - 2r$ we have

$$0 = \chi^\lambda(2r \cup \nu) = \pm[\chi^{\lambda-B}(\nu) - \chi^{\lambda-B'}(\nu)].$$

Let $\alpha = \lambda - B$, so $\alpha' = \lambda - B'$. Then for all ν we get that $\chi^\alpha(\nu) = \chi^{\alpha'}(\nu)$. Hence $\alpha = \alpha'$, which is impossible. This argument is due to to S. Sundaram (1984).

7.55. **a.** Since $\mathfrak{S}_n$ is generated by transpositions (even adjacent transpositions), we have $\rho^\lambda(\mathfrak{S}_n) \subset \mathrm{SL}(m, \mathbb{C})$ if and only if $\det \rho^\lambda(w) = 1$, where w has cycle type $\langle 21^{n-2}\rangle$. Now $\mathrm{tr}\, \rho^\lambda(w) = \chi^\lambda(21^{n-2})$, so $\rho^\lambda(w)$ has $\frac{1}{2}[f^\lambda + \chi^\lambda(21^{n-2})]$ eigenvalues equal to 1 and $\frac{1}{2}[f^\lambda - \chi^\lambda(21^{n-2})]$ eigenvalues equal to -1. Hence $\rho^\lambda(\mathfrak{S}_n) \subset \mathrm{SL}(m, \mathbb{C})$ if and only if $\frac{1}{2}[f^\lambda - \chi^\lambda(21^{n-2})]$ is even, and the proof follows from Exercise 7.51. This result is due to L. Solomon (unpublished).

7.56. **a.** If we rotate 180° an SSYT of shape θ, then we obtain a *reverse* SSYT of shape θ^r, and conversely. Now use Proposition 7.10.4.

b. Use (a) and the fact that $(B_\alpha)^r = B_{\tilde{\alpha}}$.

7.57. Let $u \in \lambda$. Let u_1 be the lowest square of λ in the same column as u, and let u_2 be the rightmost square of λ in the same row as u. Then there is a border strip B_u beginning at u_1 and ending at u_2, and this establishes a bijection between the squares and the border strips of λ. Hence the number of border strips is n. Note also that $\#B_u = h(u)$, the hook length at u.

7.58. Suppose λ has an even hook length. It is easy to see that λ then has a hook of length two. Remove it from the diagram of λ. The resulting diagram has one less even hook length and one less odd hook length than λ (see Exercise 7.59(c)), so the number of odd hook lengths minus the number of even hook lengths is unchanged. Continue removing hooks of length two until all hook lengths are

odd. The resulting partition must then be of the form $(k, k-1, \ldots, 1)$ for some k. For references, see the solution to Exercise 7.59.

7.59. The code C_λ of the partition λ was first defined by S. Comét, *Numer. Math.* **1** (1959), 90–109, and was further developed by J. B. Olsson, *Math. Scand.* **61** (1987), 223–247. A technique equivalent to the code of a partition is the theory of bead configurations and abaci, developed in G. D. James and A. Kerber, *The Representation Theory of the Symmetric Group*, Addison-Wesley, 1981 (Ch. 2.7), based on work of T. Nakayama, H. K. Farahat, G. D. James, B. Wagner, G. de B. Robinson, and D. E. Littlewood.

a.–b. These are straightforward consequences of the relevant definitions.

c. This follows easily from (a) and (b).

d. Let C_λ^j be the subsequence $\cdots c_{j-2p}c_{j-p}c_jc_{j+p}c_{j+2p}\cdots$ of C_λ. The operation described in (b) is equivalent to choosing some $0 \le j < p$, and then replacing two consecutive terms 10 in C_λ^j with 01. From this it is clear that any order of performing such operations will result in the same final sequence C_λ^* for which no further operations are possible. If $C_\lambda^* = C_\mu$ then by (b) we have that μ is the (unique) p-core of λ. For a very general approach to uniqueness results such as this exercise or Exercise 3.9(a), see K. Eriksson, Ph.D. thesis, Kungl. Tekniska Högskolan, Stockholm, 1993, and *Discrete Math.* **153** (1996), 105–122.

e. If $p = 1$ then clearly $\mu = \emptyset$ and $Y_{1,\emptyset} = Y$. Suppose that $C_\lambda^j = C_{\lambda^j}$, where C_λ^j is defined in (d). Then (d) shows that the map $\lambda \mapsto (\lambda^0, \lambda^1, \ldots, \lambda^{p-1})$, where λ has fixed p-core μ, is a bijection between $Y_{p,\mu}$ and Y^k, with

$$|\lambda| = p(|\lambda^0| + \cdots + |\lambda^{p-1}|) + |\mu|. \tag{7.208}$$

Equation (7.178) is immediate from (7.208).

The sequence $(\lambda^0, \ldots, \lambda^{p-1})$ is known as the *p-quotient* of λ. The theory of p-cores and p-quotients was originally developed by T. Nakayama, *Japan. J. Math.* **17** (1940), 165–184, 411–423; an exposition appears in G. D James and A. Kerber, *ibid.* (Ch. 2.7). An explicit statement of the isomorphism between $Y_\emptyset$ and Y^k (which trivially generalizes to $\emptyset$ replaced with any p-core μ) was first given by S. Fomin and D. W. Stanton, Rim hook lattices, *St. Petersburg Math. J.* **9** (1998), to appear (Thm. 1.2).

f. Suppose that μ is a p-core. Let x_i be the number of squares in the first column of μ whose hook length is $\equiv i \pmod p$. Then $(x_1, \ldots, x_{p-1})$ satisfies the equation in (ii), and every solution $(x_1, \ldots, x_{p-1}) \in \mathbb{N}^{p-1}$ to (ii) is obtained exactly once in this way. Thus (i) = (ii).

Now let $g(n)$ be the number in (i). Sum equation (7.178) over all p-cores μ. Since $\sum_\mu f_\mu(n) = p(n)$ (the number of partitions of n), we get

$$\prod_{i\ge1}(1-x^i)^{-1} = \left(\sum_{n\ge0} g(n)x^n\right)\cdot\prod_{i\ge1}(1-x^{pi})^{-p},$$

as desired.

g. The only partitions with no even hook length are the "staircases" $(n, n-1, \dots, 1)$. Hence we get

$$\prod_{i\geq 1} \frac{1-x^{2i}}{1-x^{2i-1}} = \sum_{n\geq 0} x^{\binom{n+1}{2}},$$

an identity due to C. F. Gauss (see e.g. [**1.1**, Cor. 2.10]).

h. The bijection $Y_\emptyset \xrightarrow{\cong} Y^k$ shows that the left-hand side of (7.179) is given by the sum $\sum_{t\in(Y^k)_n} e(t)^2$, where $(Y^k)_n$ is the set of elements of Y^k of rank n, and $e(t)$ is the number of saturated chains of Y^k between $\hat{0}$ and t. Hence

$$\begin{aligned}\sum_{\lambda\in C_p(n)} (f_p^\lambda)^2 &= \sum_{i_1+\cdots+i_p=n} \sum_{\lambda^1\vdash i_1} \cdots \sum_{\lambda^p\vdash i_p} \left[\binom{n}{i_1,\dots,i_p} f^{\lambda^1}\cdots f^{\lambda^p}\right]^2 \\ &= \sum_{i_1+\cdots+i_p=n} \binom{n}{i_1,\dots,i_p}^2 i_1!\cdots i_p! \\ &= p^n\, n! \qquad \text{[why?].}\end{aligned}$$

This exercise gives a glimpse of a body of results concerned with hook lengths divisible by p. Some references not already mentioned include J. B. Olsson, *Math. Scand.* **38** (1976), 25–42; A. N. Kirillov, A. Lascoux, B. Leclerc, and J.-Y. Thibon, *C. R. Acad. Sci. Paris, Sér. I* **318** (1994), 395–400; D. W. Stanton and D. E. White, *J. Combinatorial Theory (A)* **40** (1985), 211–247; [**96**, Exams. I.1.8-I.1.11, pp. 12–16]. (Numerous other examples in [**96**] are related.)

7.60. **a.** Let the successive squares of θ, reading from left to right and bottom to top, be $u_1, u_2, \dots, u_{rs}$. By induction on r it suffices to find a border strip λ/μ^{r-1} of λ contained in θ such that $|\lambda/\mu^{r-1}| = s$ and such that when we remove λ/μ^{r-1} from θ, the connected components thus formed (either one or two of them) will have a number of squares divisible by s.

Define $\theta_i = \{u_{is-s+1}, u_{is-s+2}, \dots, u_{is}\}$, $1 \leq i \leq r$. Let j be the least positive integer for which u_{js+1} does not lie to the right of u_{js}. The integer j exists since u_{rs+1} is undefined and hence doesn't lie to the right of u_{rs}. Since j is minimal, u_{js-s+1} lies to the right of u_{js-s}. Hence θ_j is a border strip with the desired properties. This argument, due to A. M. Garsia and R. Stanley, appears in R. Stanley, *Linear and Multilinear Algebra* **16** (1984), 3–27 (Lemma 7.3). Since the size (number of squares) of a border strip of λ is a hook length of λ (see the solution to Exercise 7.57), the second assertion of the exercise follows. One could also solve this exercise using Exercise 7.59(a)–(b).

b. Regarding s as fixed, define for any integer k

$$k^* = \begin{cases} k/s & \text{if } s \mid k \\ 0 & \text{otherwise.}\end{cases}$$

Relabel the μ_i's so that $\mu_1, \dots, \mu_j$ are not divisible by s and $\mu_{j+1}, \dots, \mu_\ell$ are, where $\ell = \ell(\mu)$. If $\chi^\lambda(\mu) \neq 0$, then by the Murnaghan–Nakayama

rule (Corollary 7.17.5) there exists a border strip tableau of shape λ and type μ. By (a), there exists a border strip tableau of shape λ and type $(\mu_1, \ldots, \mu_j, s, s, \ldots, s)$, where the number of s's is $m := \sum_i \mu_i^*$. It follows from Exercise 7.59(c) that to remove m successive border strips of size s from λ, we must have $m \leq h_s(\lambda)$, the number of hook lengths of λ divisible by s. Now the multiplicity of a primitive s-th root of unity ζ as a zero of $\prod\left(1 - q^{\mu_i}\right)$ is equal to $\#\{i : s \mid \mu_i\}$, while the multiplicity of ζ as a zero of $H_\lambda(q)$ is $h_s(\lambda)$. Clearly

$$\#\{i \;:\; s \mid \mu_i\} \leq m,$$

and the proof follows. This result first appeared in R. Stanley, *ibid.* (Cor. 7.5).

7.61. See, e.g., D. G. Duncan, *J. London Math. Soc.* **27** (1952), 235–236, or Y. M. Chen, A. M. Garsia, and J. B. Remmel, *Contemp. Math.* **34** (1984), 109–153.

In general, for any $f \in \Lambda$ we have $\langle f(x^k), s_\lambda \rangle = 0$ unless λ has an empty k-core. To see this, it suffices by linearity to assume $f = p_\mu$. Then

$$\langle p_\mu(x^k), s_\lambda \rangle = \langle p_{k\mu}, s_\lambda \rangle = \chi^\lambda(k\mu).$$

By the Murnaghan–Nakayama rule (Corollary 7.17.5), $\chi^\lambda(k\mu) = 0$ unless there exists a border strip tableau of shape λ and type $k\mu$. By Exercise 7.60(a), there then exists a border strip tableau of shape λ and type $\langle k^m \rangle$ (where $\lambda \vdash km$). Hence λ has an empty k-core.

7.62. Compute $\chi^\lambda(\mu 1^{n-k})$ by the Murnaghan–Nakayama rule (Corollary 7.17.5), choosing $\alpha = (\mu_1, \mu_2, \ldots, 1, 1, \ldots, 1)$.

7.63. **a.** It follows from the Murnaghan–Nakayama rule (Corollary 7.17.5) that

$$d_\lambda = n!\, s_\lambda|_{p_1=0,\, p_2=p_3=\cdots=1}. \tag{7.209}$$

From Proposition 7.7.4 and Theorem 7.12.1 we have

$$\sum_\lambda s_\lambda(x)s_\lambda(y) = \exp \sum_{n\geq 1} \frac{1}{n} p_n(x)p_n(y).$$

Set $p_1(y) = 0$ and $p_2(y) = p_3(y) = \cdots = 1$ to get

$$\begin{aligned}\sum_\lambda \frac{d_\lambda}{n!} s_\lambda &= \exp \sum_{n\geq 2} \frac{1}{n} p_n \\ &= e^{-h_1} \sum_{n\geq 0} h_n.\end{aligned}$$

It is now a simple matter to pick out the degree n terms to obtain the stated result.

b. It is easy to see from Pieri's rule (Theorem 7.15.7) that

$$\left\langle h_1^{n-k} h_k, s_{\langle j, 1^{n-j}\rangle} \right\rangle = \begin{cases} \binom{n-k}{j-k}, & k > 0 \\ \binom{n-1}{j-1}, & k = 0. \end{cases}$$

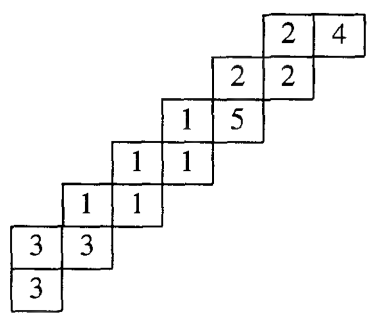
Hence from (a) we get

$$\begin{aligned} d_{(j,1^{n-j})} &= \sum_{i=1}^{j}(-1)^{n-i}(n)_i\binom{n-i}{j-i} + (-1)^n\binom{n-1}{j-1} \\ &= (-1)^{n-j}\binom{n}{j}\left[\sum_{i=0}^{j}(-1)^{j-i}\frac{j!}{(j-i)!} - (-1)^j\right] + (-1)^n\binom{n-1}{j-1} \\ &= (-1)^{n-j}\binom{n}{j}D_j + (-1)^{n-1}\binom{n-1}{j}, \end{aligned}$$

using equation (2.11). This result was first obtained (stated slightly differently, and with a different proof) by S. Okazaki, Ph.D. thesis, Massachusetts Institute of Technology; 1992 (Cor. 1.3).

7.64. a. Just read an SYT of shape τ_n from right to left and from top to bottom, to obtain an alternating permutation of $[n]$ (as defined at the end of Section 3.16). This procedure establishes a bijection between SYT of shape τ_n and alternating permutations of $[n]$. Since E_n is the number of alternating permutations of $[n]$ (as shown at the end of Section 3.16), the result follows.

b. Apply the Murnaghan–Nakayama rule (Corollary 7.17.5) to the skew shape τ_n. When n is odd, τ_n has no even-length border strips, so assume μ has $2r + 1$ odd parts. Consider a border strip tableau B of type μ, such as



so $\mu = (5, 3, 3, 1, 1)$. Reading the numbers from right to left and from top to bottom *without repetition* (e.g., here we get 4 2 5 1 3) gives an alternating permutation, and always $\mathrm{ht}(B) = k - r$. This gives the desired bijection.

c. Similar to (b), though a little more complicated.

Parts (b) and (c) are originally due to H. O. Foulkes, *Discrete Math.* **15** (1976), 311–324, who gave a more complicated proof.

Another approach was suggested by I. M. Gessel. Using e.g. the Jacobi–Trudi identity, one shows that

$$\sum_{n\geq 0}\left(\mathrm{ch}\,\chi^{\tau_n}\right)t^n = \frac{1}{\sum_{n\geq 0}(-1)^n h_{2n}t^{2n}} + \frac{\sum_{n\geq 0} h_{2n+1}t^{2n+1}}{\sum_{n\geq 0}(-1)^n h_{2n}t^{2n}}.$$

One can expand the right-hand side in terms of the p_λ's and compute coefficients explicitly.

7.65. a. First one shows that

$$\operatorname{ch}\psi_n = \sum_{k=0}^{n}(-1)^k h_1^{n-k} h_k,$$

after which it is easy to compute the character values $\psi_n(w)$. The symmetric function ch ψ_n was first considered by I. M. Gessel and C. Reutenauer, *J. Combinatorial Theory (A)* **64** (1993), 189–215 (Thm. 1, p. 206). Note that it follows from the equation deg $\psi_n = D_n$ that D_n is equal to the number of $w \in \mathfrak{S}_n$ whose largest descent has the same parity as n. This fact was first shown by J. Désarménien, in *Actes 8e Séminaire Lotharingien*, Publ. 229/S08, IRMA, Strasbourg, 1984, pp. 11–16. A generalization was given by J. Désarménien and M. L. Wachs, in *Actes 19e Séminaire Lotharingien*, Publ. 361/S19, IRMA, Strasbourg, 1988, pp. 13–21.

7.66. a. We will illustrate the proof with the example $\lambda/\mu = 8877/211$. Consider the ten zigzag dashed paths in Figure 7-23.. Each such path consists of a number of horizontal or vertical steps (or edges) from the interior of a square to the interior of an adjacent square. Every border strip of a border-strip decomposition of λ/μ cannot contain three consecutive squares that a dashed path passes through. Equivalently, let S be the set of edges e of the dashed paths with the property that the two squares through which e passes belongs to the same border strip. Then S cannot contain two consecutive edges of any dashed path. Conversely, if we choose a subset S of the edges of the dashed paths such that S contains no two consecutive edges on any dashed path, then there is a unique border-strip decomposition of λ/μ with the following property: Let e be any edge of a dashed path, and let u and v be the two squares through which e passes. Then u and v belong to the same border strip if and only if $e \in S$.

It follows that $d(\lambda/\mu)$ is equal to the number of ways to choose S. The number of ways to choose a set of edges, no two consecutive, from a path

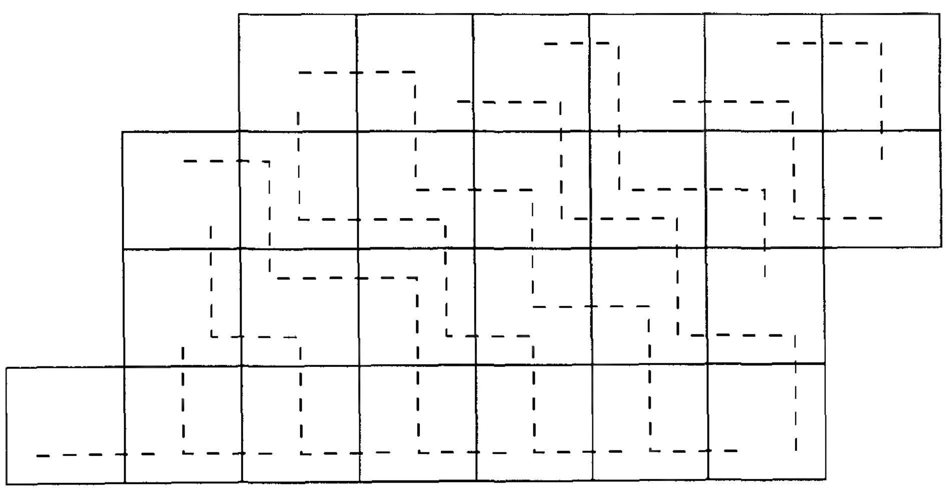


Figure 7-23. Dashed paths corresponding to the shape 8877/211.

of length n is the Fibonacci number F_{n+2} (see Exercise 1.14(a)), and the proof follows.

This exercise is actually a special case of Supplementary Exercise 3.14 from Volume 1 (second printing), which appeared in R. Stanley, Problem 10199, *Amer. Math. Monthly* **99** (1992), 162; solution by W. Y. C. Chen, **101** (1994), 278–279.

b. If P is a path of length $m-1$, then it follows from Exercise 1.13 that $\sum_T q^{\#T} = \sum_i \binom{m-i}{i} q^i$, where T ranges over all sets of edges, no two consecutive, of P. The proof is now a straightforward generalization of (a).

7.67. a. Immediate from the Murnaghan–Nakayama rule (Corollary 7.17.5).

b. Identify each C_i with the sum of its elements in the group algebra $\mathbb{C}G$. We use the standard result (e.g., [**15**, §229 and §236]) that the elements

$$E_r = \frac{d_r}{|G|}\sum_{j=1}^{t} \bar{\chi}_j^r C_j, \quad 1 \le r \le t, \tag{7.210}$$

form a complete set of orthogonal idempotents for the center of $\mathbb{C}G$. By the orthogonality of characters, inverting (7.210) yields

$$C_j = |C_j| \sum_{r=1}^{t} \frac{\chi_j^r}{d_r} E_r.$$

Since the E_i's are orthogonal idempotents (i.e., $E_r E_s = \delta_{rs} E_r$), it follows that

$$\begin{aligned} C_{i_1}\cdots C_{i_m} &= |C_{i_1}|\cdots|C_{i_m}| \sum_{r=1}^{t} \frac{\chi_{i_1}^r \cdots \chi_{i_m}^r}{d_r^m} E_r \\ &= \frac{|C_{i_1}|\cdots|C_{i_m}|}{|G|} \sum_{r=1}^{t} \frac{\chi_{i_1}^r \cdots \chi_{i_m}^r}{d_r^{m-1}} \sum_{k=1}^{t} \bar{\chi}_k^r C_k \\ &= \frac{|C_{i_1}|\cdots|C_{i_m}|}{|G|} \sum_{k=1}^{t} C_k \sum_{r=1}^{t} \frac{1}{d_r^{m-1}} \chi_{i_1}^r \cdots \chi_{i_m}^r \bar{\chi}_k^r. \end{aligned}$$

Expanding in terms of the basis G and taking the coefficient of w on both sides completes the argument. This result appears for instance in [**67**, Thm. 6.3.1] (in a somewhat more general form). Another reference is I. M. Isaacs, *Character Theory of Finite Groups*, Academic Press, New York, 1976; reprinted by Dover, New York, 1994 (Problem 3.9).

c. In (b) let $G = \mathfrak{S}_n$ and let $C_{i_1}, \ldots, C_{i_m}$ all be the conjugacy class consisting of the n-cycles. Let C_k be the conjugacy class consisting of the identity element. Then equation (7.180) reduces immediately to (7.181), using the fact that $\chi^{\langle n-s,1^s\rangle}(1^n) = f^{\langle n-s,1^s\rangle} = \binom{n-1}{s}$.

This argument appears in R. Stanley, *Discrete Math.* **37** (1981), 255–262. A survey of work related to the multiplication of conjugacy classes in $\mathfrak{S}_n$ appears in A. Goupil, *Contemp. Math.* **178** (1994), 129–143. For

some recent work not mentioned in this survey, see I. P. Goulden, *Trans. Amer. Math. Soc.* **344** (1994), 421–440; I. P. Goulden and D. M. Jackson, *J. Algebra* **166** (1994), 364–378; D. M. Jackson, *Tran. Amer. Math. Soc.* **299** (1987), 785–801; D. M. Jackson and T. I. Visentin, *Trans. Amer. Math. Soc.* **322** (1990), 353–363, 365–376; S. V. Kerov, *C. R. Acad. Sci. Paris, Sér. I* **316** (1993), 303–308 (Prop. 2.2); D. Zagier, *Nieuw Arch. Wisk. (4)* **13** (1995), 489–495; and [**96**, Exams. I.7.24–I.7.25, pp. 131–134].

d. When n is even, the terms indexed by i and $n-1-i$ cancel, so the sum is 0. Alternatively, the product of three n-cycles is an odd permutation (when n is even) and hence cannot equal the identity permutation. When n is odd, the asserted result is equivalent to the identity

$$\sum_{k=0}^{r} \frac{(-1)^k}{\binom{r}{k}} = \frac{2(r+1)}{r+2}, \qquad r \text{ even}, \tag{7.211}$$

where we have set $r = n-1$. Recall the beta function integral

$$\int_0^1 t^k(1-t)^{r-k}dt = \frac{k!\,(r-k)!}{(r+1)!}.$$

Multiply by $(-1)^k$, sum on k from 0 to n, bring the sum inside the integral, evaluate the sum explicitly, and integrate to get the stated result. This argument was suggested by D. W. Stanton. Of course equation (7.211) is not new, and there are many other ways to prove it. An independent derivation of this exercise is due to A. D. Mednykh, *Comm. Alg.* **18** (1990), 1517–1533 (eqn. (32)).

7.68. a. Note that $uvu^{-1}v^{-1} = u(vu^{-1}v^{-1})$, a product of u and a conjugate of u^{-1}. Let C and $\bar{C}$ be the conjugacy classes of G containing u and u^{-1}, respectively (so $|C| = |\bar{C}|$). If y is a fixed conjugate of u^{-1}, then there are $|G|/|C|$ elements $v \in G$ satisfying $y = vu^{-1}v^{-1}$. Hence

$$f(w) = \sum_C \frac{|G|}{|C|}\#\{(u,y) \in C \times \bar{C} \,:\, w = uy\}, \tag{7.212}$$

where C ranges over the conjugacy classes of G. Let $\chi^1, \dots, \chi^t$ denote the irreducible characters of G, and let χ^r_C denote the value of χ^r at any element of C. By Exercise 7.67(b) we have

$$\#\{(u,y) \in C \times \bar{C} \,:\, w = uy\} = \frac{|C|^2}{|G|} \sum_{r=1}^{t} \frac{1}{d_r} \chi^r_C\, \bar{\chi}^r_C\, \bar{\chi}^r(w),$$

since $|C| = |\bar{C}|$ and $\chi^r(v^{-1}) = \bar{\chi}^r(v)$. Hence

$$f = \sum_C |C| \sum_{r=1}^{t} \frac{1}{d_r} \chi^r_C\, \bar{\chi}^r_C\, \bar{\chi}^r,$$

so

$$\langle f, \chi^r \rangle = \frac{1}{d_r} \sum_C |C| \, \bar{\chi}_C^r \, \chi_C^r$$
$$= \frac{|G|}{d_r},$$

by the orthogonality of characters, and the proof follows. This result was known to many researchers in finite groups. A closely related problem appears in the book of I. M. Isaacs cited above (Problem 3.10). It is also implicit in M. Leitz, *Arch. Math. (Basel)* **67** (1996), 275–280, and some of the references given there.

b. This problem has been looked at by a number of group theorists, such as J. L. Alperin, I. M. Isaacs, and L. Solomon.

c. For any class function F on $\mathfrak{S}_n$ we have

$$\operatorname{ch} F = \sum_{\lambda \vdash n} \langle F, \chi^\lambda \rangle \, s_\lambda.$$

Now let $F = f$ (as defined in (a)).

d. For any class function g on $\mathfrak{S}_n$ and any $w \in \mathfrak{S}_n$, we have $g(w) = \langle g, p_\lambda \rangle$. Hence by equation (7.182), we have

$$f_n = \sum_{\lambda \vdash n} H_\lambda \langle s_\lambda, p_\lambda \rangle.$$

By Exercise 7.67(a) there follows

$$f_n = \sum_{k=0}^{n-1} (-1)^k H_{\langle n-k, 1^k \rangle}$$
$$= \sum_{k=0}^{n-1} (-1)^k n(n-k-1)! \, k!.$$

Now use equation (7.211). This result is equivalent to equation (43) in A. D. Mednykh, *Comm. Alg.* **18** (1990), 1517–1533.

e. Put $x_1 = \cdots = x_q = 1$ and $x_i = 0$ for $i > q$ in (7.182) and use Corollary 7.21.4.

f. It follows from (e) that

$$E_n = \frac{1}{n!} \frac{d}{dq} \sum_{\lambda \vdash n} \prod_{t \in \lambda} [q + c(t)] \bigg|_{q=1}. \tag{7.213}$$

There are three cases: (i) No content of λ is equal to -1. Then $\lambda = (n)$, and the contribution of λ to E_n in equation (7.213) is H_n.

(ii) λ has exactly one content equal to -1. Then λ has the form $\langle a, b, 1^k \rangle$, where $a \geq b > 0$, $k \geq 0$, and $a + b + k = n$. In this case the contribution of λ to E_n is

$$\frac{1}{n!} \prod_{\substack{t \in \lambda \\ t \neq (2,1)}} [1 + c(t)] = (-1)^k a! \, (b-1)! \, k!. \tag{7.214}$$

(iii) λ has more than one content equal to -1. Then the contribution of λ to E_n is 0.

When we sum (7.214) over all (a, b, k) satisfying $a \geq b > 0, k \geq 0$, and $a + b + k = n$, then it is not hard to see (using equation (7.211)) that we get the right-hand side of (7.183) except for the term H_n, which already arose from $\lambda = (n)$.

g. Let Γ_j be the functional on symmetric functions defined by

$$\Gamma_j(f) = \frac{\partial}{\partial p_j} f \bigg|_{p_i=1},$$

where $g|_{p_i=1}$ indicates that we are to expand g as a polynomial in the p_i's and then set each $p_i = 1$. Thus from (7.182) we have

$$e_{nj} = \frac{1}{n!} \Gamma_j \left(\sum_{\lambda \vdash n} H_\lambda s_\lambda \right).$$

Let $m_j(\mu)$ denote the number of parts of μ equal to j, and note that

$$\begin{aligned} \Gamma_j(p_\mu) &= m_j(\mu) \\ &= \left\langle p_\mu, \sum_\lambda z_\lambda^{-1} m_j(\mu) p_\mu \right\rangle \\ &= \left\langle p_\mu, p_j \frac{\partial}{\partial p_j} h_n \right\rangle. \end{aligned}$$

But from $\sum_n h_n = \exp\left(\sum_n p_n/n\right)$ there follows

$$\frac{\partial}{\partial p_j} h_n = \frac{1}{j} h_{n-j}.$$

Hence from the linearity of Γ_j we get that for *any* $f \in \Lambda^n$,

$$\Gamma_j(f) = \left\langle f, \frac{1}{j} p_j h_{n-j} \right\rangle.$$

By Theorem 7.17.1 we have

$$p_j h_{n-j} = \sum_\rho (-1)^{\mathrm{ht}(\rho/(n-j))} s_\rho,$$

summed over all partitions $\rho \supseteq (n - j)$ for which $\rho/(n - j)$ is a border strip B of size j. For each $-1 \leq i \leq j - 1$ there is exactly one such border strip B_i with $i + 1$ squares in the first column, except that B_{2j-n-1} does not exist when $2j > n$. Write ρ^i for the partition ρ for which $\rho/(n - j) = B_i$. It follows that for $\lambda \vdash n$,

$$\Gamma_j(s_\lambda) = \begin{cases} \dfrac{1}{j}(-1)^{\mathrm{ht}(B_i)} & \text{if } \lambda = \rho^i \\ 0, & \text{otherwise.} \end{cases}$$

(This formula can also be obtained by showing that $\partial/\partial p_j$ is adjoint to

multiplication by $\frac{1}{j}p_j$, and that $s_\lambda|_{p_i=1}=0$ unless $\lambda=(n)$.) In particular,

$$\begin{aligned} e_{nj} &= \frac{1}{n!}\Gamma_j\left(\sum_{\lambda\vdash n} H_\lambda s_\lambda\right) \\ &= \frac{1}{n!\,j}\sum_{i=-1}^{j-1}{}'(-1)^{\mathrm{ht}(B_i)}H_{\rho^i}. \end{aligned}$$

When $i=-1$ we have $\rho^{-1}=(n)$ and

$$\frac{1}{n!}(-1)^{\mathrm{ht}(B_{-1})}H_{\rho^{-1}}=1.$$

When $0\le i\le j-1$ (omitting $i=2j-n-1$ when $2j>n$) it is not hard to check (considering separately the cases $2j\le n+i$ and $2j>n+i+1$) that

$$\frac{1}{n!}(-1)^{\mathrm{ht}(B_i)}H_{\rho^i}=\frac{(-1)^i(n-j+i+1)}{\binom{n}{j}\binom{j-1}{i}(n-2j+i+1)},$$

and the proof follows. This result is due to R. Stanley and J. R. Stembridge (unpublished). For the problem of computing the expected number of j-cycles of more general expressions than $uvu^{-1}v^{-1}$, see A. Nica, *Random Structures and Algorithms* **5** (1994), 703–730.

7.69. a. The square of a cycle of odd length n is an n-cycle, while the square of a cycle of even length n is the product of two cycles of length $n/2$. It then follows from the exponential formula (Corollary 5.1.9) that

$$\begin{aligned} \sum_{n\ge 0}\frac{1}{n!}\sum_{w\in\mathfrak{S}_n}p_{\rho(w^2)} &= \exp\left(\sum_{\substack{n\ge 1\\ n\text{ odd}}}\frac{1}{n}p_n+\sum_{\substack{n\ge 1\\ n\text{ even}}}\frac{1}{n}p_{n/2}^2\right) \\ &= \exp\sum_{n\ge 1}\frac{1}{n}\left(\sum_i x_i^n+\sum_{i<j}(x_ix_j)^n\right) \\ &= \frac{1}{\prod_i(1-x_i)\cdot\prod_{i<j}(1-x_ix_j)} \\ &= \sum_{n\ge 0}\sum_{\lambda\vdash n}s_\lambda, \end{aligned}$$

the last step by Corollary 7.13.8, so we get

$$\frac{1}{n!}\sum_{w\in\mathfrak{S}_n}p_{\rho(w^2)}=\sum_{\lambda\vdash n}s_\lambda.$$

b. Let $f(w)=\sum_{\lambda\vdash n}\chi^\lambda(w)$, so $\mathrm{ch}\,f=\sum_\lambda s_\lambda$. We need to show that $\mathrm{ch}\,f$ is *p-positive*, i.e., a nonnegative linear combination of p_μ's. Now use (a). This argument shows in fact that

$$\sum_{\lambda\vdash n}\chi^\lambda(w)=\#\{u\in\mathfrak{S}_n : w=u^2\}.$$

See also [**96**, Exam. 11, p. 120]. More generally, it follows from the work of Frobenius and Schur (see Isaacs, *ibid.* (Ch. 4) for an exposition) that if G is a finite group for which every complex representation is equivalent to a real representation, then for any $w \in G$ we have

$$\sum_{\chi \in \hat{G}} \chi(w) = \#\{u \in G \,:\, w = u^2\},$$

where $\hat{G}$ denotes the set of irreducible characters of G.

c. This is a result of T. Scharf, *Bayreuther Math. Schr.*, No. 38 (1991), 99–207, and *J. Algebra* **139** (1991), 446–457. (See also [**67**, §6.2].) Scharf shows the following. Let $\mathrm{Par}_k(n)$ be the set of all partitions of n all of whose parts divide k. For each $\lambda \in \mathrm{Par}_k(n)$, choose an element $w_\lambda \in \mathfrak{S}_n$ of cycle type λ. Let $\zeta = e^{2\pi i/k}$. Define a one-dimensional character ψ^λ on the cyclic group generated by w_λ by

$$\psi^\lambda(w_\lambda) = \zeta^{k/\lambda_1 + \cdots + k/\lambda_\ell},$$

where $\ell = \ell(\lambda)$. This character extends naturally to a one-dimensional character ψ^λ of the centralizer $C(w_\lambda)$ of w_λ. Then

$$r_k = \sum_{\lambda \in \mathrm{Par}_k(n)} \mathrm{ind}_{C(w_\lambda)}^{\mathfrak{S}_n} \psi^\lambda.$$

Hence r_k is a character of $\mathfrak{S}_n$, and the proof follows.

A proof based more on the theory of symmetric functions was given by J.-Y. Thibon, *Bayreuther Math. Schrift.*, no. 40 (1992), 177–201 (Cor. 5.2). We give a somewhat simplified version of this proof (generalizing the argument in (a)) as follows. Let $\theta_{k,n} = \mathrm{ch}\, r_{n,k}$. Since the k-th power of an n-cycle is a product of (n,k) cycles of length $n/(n,k)$ (where (n,k) denotes the g.c.d. of n and k), it follows immediately from the exponential formula (Corollary 5.1.9) that

$$\sum_{n \geq 0} \theta_{k,n} = \exp \sum_{n \geq 1} \frac{1}{n} p_{n/(n,k)}^{(n,k)}. \tag{7.215}$$

Let L_d be the symmetric function of equation (7.191). A simple inclusion–exclusion argument shows that

$$\sum_{n \geq 1} \frac{1}{n} p_{n/(n,k)}^{(n,k)} = \sum_{d|k} \sum_{n \geq 1} \frac{1}{n} L_d(x^n), \tag{7.216}$$

where $L_d(x^n) = L_d(x_1^n, x_2^n, \ldots)$. Equation (7.215) is then equivalent to the plethystic formula

$$\sum_{n \geq 0} \theta_{k,n} = \sum_{n \geq 0} h_n \left[\sum_{d|k} L_d \right]. \tag{7.217}$$

Equivalently, if $h = \sum_{n \geq 0} h_n$, then

$$\sum_{n \geq 0} \theta_{k,n} = \prod_{d|k} h[L_d].$$

Now h_n is just the Schur function s_n, while L_d is a nonnegative (integer) linear combination of Schur functions by Exercise 7.89(b). Hence by Theorem A2.5 of Appendix 2, the right-hand side of (7.217) is also a nonnegative linear combination of Schur functions, and the proof follows. We don't know whether r_k (extended in an obvious way to any finite group) is a character of any finite group G for which every representation can be realized over $\mathbb{Z}$. (The quaternion group of order eight shows that it does not suffice just to assume that G is an IC-group, as defined in (j).)

d. It suffices by iteration to assume that $m = 2$ (though the general case can also be proved directly). Let $a = f_1$ and $b = f_2$. We follow the notation of Exercise 7.67(b). We have

$$h(w) = \sum_{i=1}^{t}\sum_{j=1}^{t} a_i b_j \,\#\{(u, v) \in G \times G \,:\, u \in C_i,\ v \in C_j,\ uv = w\}.$$

By Exercise 7.67(b), we have

$$\begin{aligned}\#\{(u, v) &\in G \times G \,:\, u \in C_i,\ v \in C_j,\ uv = w\}\\ &= \frac{|C_i| \cdot |C_j|}{|G|} \sum_{r=1}^{t} \frac{1}{d_r} \chi_i^r \chi_j^r \bar{\chi}^r(w).\end{aligned}$$

Hence

$$\begin{aligned}\langle h, \chi^r\rangle &= \frac{1}{|G|}\sum_{i=1}^{t}\sum_{j=1}^{t} |C_i| \cdot |C_j| a_i b_j \frac{1}{d_r} \chi_i^r \chi_j^r\\ &= \frac{|G|}{d_r}\left(\frac{1}{|G|}\sum_{i=1}^{t} |C_i| a_i \chi_i^r\right)\left(\frac{1}{|G|}\sum_{j=1}^{t} |C_j| b_j \chi_j^r\right)\\ &= \frac{|G|}{d_r}\langle f, \chi^r\rangle \cdot \langle g, \chi^r\rangle,\end{aligned}$$

and the proof is complete. It is possible to view this result as a special case of the theorem that the Fourier transform converts convolution to multiplication.

e. Define $\psi^\lambda : \mathfrak{S}_n \to \mathbb{Z}$ by $\psi^\lambda(w) = 1$ if $\rho(w) = \lambda$, and $\psi^\lambda(w) = 0$ otherwise. It is then easy to check that

$$\mathrm{ch}\, F_{\psi^\lambda, \psi^\mu} = (\mathrm{ch}\, \psi^\lambda) \square \left(\mathrm{ch}\, \psi^\mu\right).$$

By bilinearity we get $\mathrm{ch}\, F_{f,g} = (\mathrm{ch}\, f) \square (\mathrm{ch}\, g)$ for any class functions f, g on $\mathfrak{S}_n$. Now put $f = \chi^\lambda$, $g = \chi^\mu$, and use (7.184) to deduce (7.185). The steps can be reversed to deduce (7.184) from (7.185).

f. Letting r_k be as in (c), we have

$$h(w) = \prod_{u_1 \cdots u_m = w} r_{a_1}(u_1) \cdots r_{a_m}(u_m).$$

The proof now follows from parts (c) and (f).

Some results closely related to (d) and (f) appear in A. Kerber and B. Wagner, *Arch. Math. (Basel)* **35** (1980), 252–262. Indeed, if one uses the

fact (Isaacs, *ibid.*, Lemma 4.4, p. 49) that for any finite group G and any irreducible character χ we have

$$\langle r_k, \chi \rangle = \frac{1}{|G|} \sum_{w \in G} \chi(w^k),$$

then Satz 1 of Kerber and Wagner is equivalent to our (d) when $f_i = r_{a_i}$. Another related paper (though not as closely) is L. Solomon, *Arch. Math. (Basel)* **20** (1969), 241–247.

g. Let $a = uvu$ and $b = uv$. Then $u = b^{-1}a$ and $v = a^{-1}b^2$, so as (u, v) ranges over $G \times G$, so does (a, b). But $uvu^2vuv = ab^2$, which is clearly equidistributed over G. Thus we get

$$\#\{(u, v) \in G \times G \,:\, w = uvu^2vuv\} = |G|.$$

The main point is that the substitution $u = b^{-1}a$ and $v = a^{-1}b^2$ is invertible because the homomorphism $\varphi : F_2 \to F_2$ (where F_2 is the free group on generators x, y) defined by $\varphi(x) = y^{-1}x$ and $\varphi(y) = x^{-1}y^2$ is an automorphism of F_2. For the classification of automorphisms of free groups, see e.g. M. Hall, Jr., *The Theory of Groups*, Macmillan, New York, 1959 (Thm. 7.3.4).

h. We get from Exercise 7.68(a) and parts (a) and (d) of this exercise that $\langle f, \chi^\lambda \rangle = \langle g, \chi^\lambda \rangle = H_\lambda$ for all $\lambda \vdash n$, so the proof follows.

For a bijective proof, note that every element in $\mathfrak{S}_n$ is conjugate to its inverse. Hence we may replace the equation $w = u(vu^{-1}v^{-1})$ with $w = u(vuv^{-1})$ and maintain a simple bijection between the solutions to the two equations. Now let $a = uv$ and $v = b^{-1}$. Since $u = ab$ and $v = b^{-1}$, the elements a and b range over $\mathfrak{S}_n$ as u and v do. Hence we may replace the equation $w = uvuv^{-1}$ with $w = (ab)b^{-1}(ab)b = a^2b^2$, and the result follows. This argument is valid in any (finite) group for which every element is conjugate to its inverse, or equivalently (since $\chi(w^{-1}) = \bar{\chi}(w)$) for which every character is real, such as a (finite) Coxeter group. The case $w = 1$ was treated in Exercise 5.12 for any finite group G.

i. For $\gamma = xy^kxy^{-k}$ the second argument in (h) generalizes easily. Namely, let $a = xy^k$ and $b = y^{-1}$ to get $f_{\gamma, \mathfrak{S}_n} = f_{a^2b^{2k}, \mathfrak{S}_n}$, which is a character by (f). For $\gamma = xy^kx^{-1}y^{-k}$, note that as in (7.212) we have

$$f_{\gamma, G}(w) = \sum_C r_k(C) \frac{|G|}{|C|} \#\{(u, y) \in C \times \bar{C} \,:\, w = uy\},$$

where r_k is defined in (c). Reasoning exactly as in the solution to Exercise 7.68(a) shows that

$$\langle f, \chi \rangle = \frac{|G|}{\chi(1)} \langle r_k, \chi\bar{\chi} \rangle,$$

for every irreducible character χ of G. Since r_k is a character when $G = \mathfrak{S}_n$ by (c), it follows that f is also a character. Similar reasoning shows that if β is any word in the letters $x_1, \ldots, x_r$ and if x is a letter different from the

x_i's, then

$$\langle f_{x\beta x^{-1}\beta^{-1}}, \chi\rangle = \frac{|G|}{\chi(1)}\langle f_\beta, \chi\bar{\chi}\rangle$$

$$\langle f_{x\beta x^{-1}\beta}, \chi\rangle = \frac{|G|}{\chi(1)}\langle f_\beta, \chi^2\rangle.$$

j. When $r = 1$ (so $f_{\gamma,G} = r_k$ for some $k \in G$), it follows from the work of Frobenius, *Sitz. Köngl. Preuß. Akad. Wissen. Berlin* (1907), 428–437; *Ges. Abh.*, vol. III, Art. 78, pp. 394–403, that r_k is a difference of two characters. A sketched proof appears in Isaacs, *ibid.* (Problem 4.7). When $r > 1$, then it suffices to show (since $f_{\gamma,G}$ is a class function) that for fixed i,

$$\#\{(u_1, \dots, u_r) \in G^r : \gamma(u_1, \dots, u_r) \in C_i\} \equiv 0 \pmod{|G|}.$$

But this is exactly the special case $m = 1$ of Theorem 1 of L. Solomon, *Arch. Math. (Basel)* **20** (1969), 241–247. For a closely related result, see I. M. Isaacs, *Canad. J. Math.* **22** (1970), 1040–1046 (Thm. B).

k. If $\gamma = x^2y^2x^2y^2, x^2y^3x^2y^{-3}$, or $x^2y^2x^2y^3$, then we don't know whether $f_{\gamma,\mathfrak{S}_n}$ is a character for all n. (The case $x^2y^2x^2y^2$ has been checked for $n \le 16$, and the other two for $n \le 7$.) On the other hand, if $\gamma = x\,y^{-1}x^2y$, $x^2y^3x^{-2}y^{-3}$, or $x^2y^3x^5y^4$, then $f_{\gamma,\mathfrak{S}_n}$ is *not* a character for all n. Note also that for a word like $\gamma = x_1^5x_2^3x_3^5x_4^4x_5^5x_6^3x_7^4x_8^5$ (where every exponent occurs an even number of times), it follows from (d) that $f_{\gamma,G}$ is a character for all finite groups G.

7.70. Let $\mu^1, \dots, \mu^k \vdash n$. By Corollary 7.17.5, when the left-hand side of equation (7.186) is expanded in terms of power sums, the coefficient Q of $p_{\mu^1}(x^{(1)})\cdots p_{\mu^k}(x^{(k)})$ is given by

$$Q = \sum_{\lambda\vdash n} H_\lambda^{k-2}\prod_{i=1}^k z_{\mu^i}^{-1}\chi^\lambda(\mu^i).$$

Let C_μ denote the conjugacy class of $\mathfrak{S}_n$ consisting of permutations of cycle type μ. Since $|C_\mu| = n!/z_\mu$ (by equation (7.18)), $f^\lambda = n!/H_\lambda$ (by Corollary 7.21.6), and $\chi^\lambda(1^n) = f^\lambda$ (equation (7.79)), we have

$$Q = \left(\frac{1}{n!}\prod_{i=1}^k |C_{\mu^i}|\right)\sum_{\lambda\vdash n}\frac{1}{(f^\lambda)^{k-1}}\left(\prod_{i=1}^k\chi^\lambda(\mu^i)\right)\chi^\lambda(1^n).$$

Comparing with equation (7.180) (and using the fact that the χ^λ's are the irreducible characters of $\mathfrak{S}_n$ (Theorem 7.18.5) and that they are real), we see that

$$Q = \frac{1}{n!}\#\{(w_1, \dots, w_k) \in \mathfrak{S}_n^k : w_1\cdots w_k = \mathrm{id}\},$$

as desired.

The case $k = 0$ is equivalent to Corollary 7.12.6, while the case $k = 1$ is equivalent to Corollary 7.12.5.

Equation (7.186) first appeared in P. J. Hanlon, R. Stanley, and J. R. Stembridge, *Contemporary Math.* **138** (1992), 151–174 (Prop. 2.2), in connection

with the distribution of the eigenvalues of the matrix $AUBU^*$, where A and B are fixed $n \times n$ Hermitian matrices, and U is a random $n \times n$ matrix whose entries are independent standard complex normal random variables.

7.71. a. By the orthogonality of characters,

$$\sum_{\chi} \chi\bar{\chi}(w) = \#C(w),$$

the order of the centralizer $C(w)$ of w (the number of $v \in G$ commuting with w). This is just the number of fixed points of the action of w on G by conjugation (the number of $v \in G$ such that $wvw^{-1} = v$). Since the number of fixed points of an element w acting on a set is its character value, (i) and (ii) agree.

b. Let $\hat{G}$ denote the set of irreducible characters of G. Then

$$\begin{aligned}\langle \psi_G, \chi \rangle &= \frac{1}{\#G} \sum_{\theta \in \hat{G}} \sum_{w \in G} \theta(w)\bar{\theta}(w)\chi(w)\\ &= \sum_{w \in G} \chi(w)\left(\frac{1}{\#G}\sum_{\theta} \theta(w)\bar{\theta}(w)\right)\\ &= \sum_{w\in G} \chi(w)[G : C(w)]^{-1}.\end{aligned}$$

But $[G : C(w)]$ is the number of conjugates of w. Thus the previous sum becomes $\sum_K \chi(K)$.

c. We have

$$\begin{aligned}\operatorname{ch} \psi_n &= \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} \psi_n(w) p_w\\ &= \sum_{w\in\mathfrak{S}_n} [\mathfrak{S}_n : C(w)]^{-1} p_w\\ &= \sum_{\lambda \vdash n} p_\lambda.\end{aligned}$$

d. See A. Frumkin, *Israel J. Math. (1)* **55** (1986), 121–128; T. Scharf, *Arch. Math. (Basel)* **54** (1990), 427–429; and Y. Roichman, *Israel J. Math.* **97** (1997), 305–316.

e. For some related work, see H. Décoste, *Séries Indicatrices d'Espèces Pondérées et q-Analogues*, Publications du LACIM, vol. 2, Université du Québec à Montréal, 1989 (Example 3.7).

7.72. Let $\Lambda^k A$ denote the action of A on $\Lambda^k V$. If A has eigenvalues $\theta_1, \dots, \theta_n$, then $\Lambda^k A$ has eigenvalues $\theta_{i_1} \cdots \theta_{i_k}$, $1 \le i_1 < \cdots < i_k \le n$. Hence

$$\sum_{k=0}^{n} \left(\operatorname{tr} \Lambda^k A\right)(-q)^k = (1 - \theta_1 q)\cdots(1-\theta_n q) = \det(I - qA). \qquad (7.218)$$

Now if $w \in \mathfrak{S}_n$ has cycle type $\mu = (\mu_1, \mu_2, \dots)$ with $\ell(\mu) = \ell$, then $\det(I - qw) = \left(1 - q^{\mu_1}\right)\left(1 - q^{\mu_2}\right)\cdots\left(1 - q^{\mu_\ell}\right)$. Hence writing Ψ_k for the

character of $\mathfrak{S}_n$ acting on $\Lambda^k V$, we have

$$\sum_{k=0}^{n} \Psi_k(w)(-q)^k = \left(1 - q^{\mu_1}\right)\left(1 - q^{\mu_2}\right) \cdots \left(1 - q^{\mu_\ell}\right), \qquad (7.219)$$

so

$$\sum_{k=0}^{n} (\operatorname{ch} \Psi_k)(-q)^k = \frac{1}{n!} \sum_{\substack{w \in \mathfrak{S}_n \\ \rho(w)=\mu}} \left(1 - q^{\mu_1}\right) \cdots \left(1 - q^{\mu_\ell}\right) p_{\mu_1} \cdots p_{\mu_\ell}.$$

Note that since $p_j(-qx) = (-1)^j q^j p_j(x)$ and $\omega(p_j) = (-1)^{j-1} p_j$, we have

$$(1 - q^j) p_j(x) = \omega_y p_k(x, y)|_{y=-qx},$$

where ω_y denotes ω acting on the y-variables only. Hence

$$\begin{aligned}
\sum_{k=0}^{n} (\operatorname{ch} \Psi_k)(-q)^k &= \omega_y \left[\frac{1}{n!} \sum_{w} p_{\rho(w)}(x, y) \right]_{y=-qx} \\
&= \omega_y h_n(x, y)|_{y=-qx} \qquad \text{(by Proposition 7.7.6)} \\
&= \omega_y \sum_{j=0}^{n} s_j(x) s_{n-j}(y)|_{y=-qx} \qquad \text{(by (7.66))} \\
&= \sum_{j=0}^{n} s_j(x) s_{1^{n-j}}(y)|_{y=-qx} \qquad \text{(by Theorem 7.14.5)} \\
&= \sum_{j=0}^{n} (-q)^{n-j} s_j(x) s_{1^{n-j}}(x).
\end{aligned}$$

It is a simple matter to multiply s_j by $s_{1^{n-j}}$ by Pieri's rule (Theorem 7.15.7) and collect terms to get

$$\sum_{k=0}^{n} (\operatorname{ch} \Psi_k) q^k = \sum_{k} (s_{\lambda^k} + s_{\lambda^{k-1}}) q^k,$$

whence $\operatorname{ch} \Psi_k = s_{\lambda^k} + s_{\lambda^{k-1}}$ and $\Psi_k = \chi^{\lambda^k} + \chi^{\lambda^{k-1}}$.

Note that this result is equivalent to $\Lambda^k \chi^{(n-1,1)} = \chi^{(n-k,1^k)}$. The result of this exercise is due to A. C. Aitken, *Proc. Edinburgh Math. Soc. (2)* **7** (1946), 196–203.

7.73. The argument is similar to that of Exercise 7.72. Let $S^k A$ denote the action of A on $S^k V^*$, the space of homogeneous forms of degree k in the variables $v_1, \ldots, v_n$. Analogously to (7.218) and (7.219) we get

$$\sum_{k \geq 0} \left(\operatorname{tr} S^k A\right) q^k = \frac{1}{(1 - \theta_1 q) \cdots (1 - \theta_n q)} = \frac{1}{\det(I - qA)}$$

and

$$\sum_{k \geq 0} \psi^k(w) q^k = \frac{1}{\left(1 - q^{\mu_1}\right)\left(1 - q^{\mu_2}\right) \cdots \left(1 - q^{\mu_\ell}\right)}.$$

Hence

$$\sum_{k\geq 0}\left(\operatorname{ch}\psi^k\right)q^k = \frac{1}{n!}\sum_{\substack{w\in\mathfrak{S}_n\\ \rho(w)=\mu}}\frac{p_{\mu_1}\cdots p_{\mu_\ell}}{\left(1-q^{\mu_1}\right)\cdots\left(1-q^{\mu_\ell}\right)}$$

$$= \frac{1}{n!}\sum_{\substack{w\in\mathfrak{S}_n\\ \rho(w)=\mu}} p_\mu(1,q,q^2,\ldots)p_\mu$$

$$= \sum_{\lambda\vdash n} s_\lambda(1,q,q^2,\ldots)s_\lambda,$$

by Proposition 7.7.4 and Theorem 7.12.1.

This result is due to A. C. Aitken, *Proc. Edinburgh Math. Soc. (2)* **5** (1937), 1–13 (Thm. 2). For a "modern" λ-ring proof of this exercise and the previous one (Exercise 7.72), see J.-Y. Thibon, *Bayreuth Math. Schr.*, No. 40 (1992), 177–201 (§3).

7.74. Let the permutation $w\in\mathfrak{S}_n$, regarded as an element of $\mathrm{GL}(n,\mathbb{C})$, have eigenvalues $\theta_1,\ldots,\theta_n$. Then $\xi^\lambda(w)=s_\lambda(\theta_1,\ldots,\theta_n)$. Let $\rho(w)=(\rho_1(w), \rho_2(w),\ldots)$. Since

$$\prod_{i=1}^n(1-\theta_i q)=\prod_j(1-q^j)^{\rho_j(w)},$$

we get by the Cauchy identity (Theorem 7.12.1) that

$$\sum_\lambda \xi^\lambda(w)s_\lambda(y)=\frac{1}{\prod_{i,j}\left(1-y_i^j\right)^{\rho_j(w)}}.$$

Taking the characteristic of both sides yields

$$\sum_\lambda(\operatorname{ch}\xi^\lambda)(x)s_\lambda(y)=\frac{1}{n!}\sum_{w\in\mathfrak{S}_n}\frac{1}{\prod_{i,j}\left(1-y_i^j\right)^{\rho_j(w)}}\cdot p_{\rho(w)}(x)$$

$$=\frac{1}{n!}\sum_{w\in\mathfrak{S}_n}p_{\rho(w)}(x)p_{\rho(w)}[h](y)$$

$$=\sum_{\mu\vdash n}s_\mu(x)s_\mu[h](y),$$

and the proof follows.

Equation (7.187) appears in T. Scharf and J.-Y. Thibon, *Advances in Math.* **104** (1994), 30–58 (Thm. 5.1). The correspondence $s_\lambda\mapsto\operatorname{ch}\xi^\lambda$ is a special case of the operation of *inner plethysm*, defined as follows. Let $\sigma:\mathfrak{S}_n\to\mathrm{GL}(N,\mathbb{C})$ be any (finite-dimensional) representation of $\mathfrak{S}_n$, and let $\varphi:\mathrm{GL}(N,\mathbb{C})\to\mathrm{GL}(M,\mathbb{C})$ be a polynomial representation of $\mathrm{GL}(N,\mathbb{C})$. The composition $\varphi\sigma$ is then a representation of $\mathfrak{S}_n$, and we define the inner plethysm $f\odot g$ of the symmetric functions $f=\operatorname{ch}\sigma$ and $g=\operatorname{char}\varphi$ by

$$f\odot g=\operatorname{ch}\varphi\sigma.$$

In particular,

$$\left(s_n + s_{(n-1,1)}\right) \odot s_\lambda = \mathrm{ch}\,\xi^\lambda,$$

since the character of the defining representation $\mathfrak{S}_n \to \mathrm{GL}(n, \mathbb{C})$ of $\mathfrak{S}_n$ is given by $\chi^n + \chi^{(n-1,1)}$. Inner plethysm was introduced by D. E. Littlewood, *Canad. J. Math.* **10** (1958), 1–16, 17–32. For further information and references, see T. Scharf and J.-Y. Thibon, *ibid.*

*

7.75. **a.** An orbit $\mathcal{O}_\mu$ of the action of $\mathfrak{S}_k$ on $\left(\binom{M}{j}\right)$ is specified by a partition $\mu = \langle 1^{m_1} 2^{m_2} \cdots n^{m_n} \rangle \vdash j$, where $\sum m_i = k$ and $\ell(\mu) = \sum m_i \leq k$. Here $\mathcal{O}_\mu$ consists of those submultisets $N \in \left(\binom{M}{j}\right)$ with m_i elements of multiplicity i. For instance, the orbit containing $\{1, 1, 1, 1, 3, 4, 4, 5, 5, 5, 5, 7, 8\}$ corresponds to the partition $\mu = \langle 1^3, 2^1, 4^2 \rangle \vdash 13$. The characteristic $\mathrm{ch}(\mathcal{O}_\mu)$ of the action of $\mathfrak{S}_k$ on $\mathcal{O}_\mu$ is just $h_{m_1} h_{m_2} \cdots h_{m_n} h_{k-\ell(\mu)}$ [why?]. Hence (setting $r = k - \ell(\mu)$),

$$\sum_j \mathrm{ch}\left(\binom{M}{j}\right) q^j = \left. \sum_{\substack{m_1, \ldots, m_n \geq 0 \\ r \geq 0}} h_{m_1} \cdots h_{m_n} h_r q^{\sum i m_i} \right|_k,$$

where $f|_k$ denotes the degree k part of the symmetric function f. We get

$$\begin{aligned} \sum_j \mathrm{ch}\left(\binom{M}{j}\right) q^j &= (1 + h_1 + h_2 + \cdots) \prod_{i=1}^n (1 + q^i h_1 + q^{2i} h_2 + \cdots)|_k \\ &= \left. \frac{1}{\prod (1 - x_i)(1 - q x_i) \cdots (1 - q^n x_i)} \right|_k \\ &= \sum_{\lambda \vdash k} s_\lambda(1, q, \ldots, q^n) s_\lambda(x) \qquad \text{(by (7.44))}, \end{aligned}$$

and the proof follows.

b. It is easy to see that U_j commutes with the action of $\mathfrak{S}_k$. A proof of injectivity for $j < kn/2$ involving only elementary linear algebra is a special case of the argument given in §6 of R. A. Proctor, M. E. Saks, and D. G. Sturtevant, *Discrete Math.* **30** (1980), 173–180. For some related work see R. A. Proctor, *J. Combinatorial Theory (A)* **54** (1990), 235–247 (especially Cor. 1).

c. We omit the easy proof that $a_j = a_{kn-j}$. Let $j < \lfloor kn/2 \rfloor$. Since U_j commutes with the action of $\mathfrak{S}_k$ and is injective, it is an injective map of $\mathfrak{S}_k$-modules. Thus every irreducible representation of $\mathfrak{S}_k$ occurs at least as often in $\mathbb{Q}\left(\binom{M}{j+1}\right)$ as in $\mathbb{Q}\left(\binom{M}{j}\right)$, so by (a) we get $a_j \leq a_{j+1}$.

The unimodality of $s_\lambda(1, q, \ldots, q^n)$ was first proved (though not stated explicitly in terms of Schur functions) by E. B. Dynkin, *Dokl. Akad. Nauk SSSR (N.S.)* **71** (1950), 221–224, and *Amer. Math. Soc. Translations, Series 2* **6** (1957), 245–378 (p. 332) (translated from *Trudy Moskov. Mat. Obšč.* **1** (1957), 39–156), in the context of the representation theory of semisimple Lie groups. For a statement of Dynkin's result avoiding the language of representation theory, see R. Stanley, in *Young Day Proceedings* (T. V. Narayana, R. M. Mathsen, and J. G. Williams, eds.), Dekker, New

York/Basel, 1980, pp. 127–136. An elegant version of Dynkin's proof, in the special case we are considering here, is given in [**96**, Exam. I.8.4, pp. 137–138]. It is based on Theorem A2.5 of Appendix 2. A proof similar to the one given here (but using wreath products instead of multisets) appears in A. Kerber and K.-J. Thürlings, *Bayreuther Math. Schr.* **21** (1986), 156–278 (Satz 2.2). For further information see R. Stanley, *Ann. New York Acad. Sci.* **576**, 1989, pp. 500–535. No proof of the unimodality of $s_\lambda(1, q, \ldots, q^n)$ is known that does not involve representation theory (or even more sophisticated tools), but see (d) below for a special case for which simpler proofs are known.

d. Apply (c) to $s_{1^k}(1, q, \ldots, q^{n-1})$ or $s_k(1, q, \ldots, q^{n-k+1})$, where these specializations are evaluated in Proposition 7.8.3. This result goes back to J. J. Sylvester, *Phil. Mag.* **5** (1878), 178–188; in *Collected Math. Papers*, vol. 3, Chelsea, New York, 1973, pp. 117–126. A number of other proofs have subsequently been given, as discussed in R. Stanley, *ibid.* In particular, a long-sought-for combinatorial proof was found by K. M. O'Hara, *J. Combinatorial Theory (A)* **53** (1990), 29–52; an exposition was given by D. Zeilberger, *Amer. Math. Monthly* **96** (1989), 590–602.

7.76. a. Let f_w denote the number of fixed points of $w \in G$ acting on T, so $f_w = \chi(w)$. Thus

$$\langle \chi, \chi \rangle = \frac{1}{\#G} \sum_{w \in G} f_w^2.$$

On the other hand, the number of fixed points of w acting on $T \times T$ is just f_w^2. Thus by Burnside's lemma (Lemma 7.24.5), the above sum is the rank of G.

b. Let χ_α denote the character of this action, so by Corollary 7.18.3 we have $\operatorname{ch} \chi_\alpha = h_\alpha$. Since ch is an isometry (Proposition 7.18.1) the rank of $\mathfrak{S}_n$ acting on $\mathfrak{S}_n/\mathfrak{S}_\alpha$ is given by $\langle h_\alpha, h_\alpha \rangle$, which by equation (7.31) is the number of $\mathbb{N}$-matrices A with $\operatorname{row}(A) = \operatorname{col}(A) = \alpha$.

c. The left cosets of $\mathfrak{S}_\alpha$ are indexed in a natural way by permutations of the multiset $M_\alpha = \{1^{\alpha_1}, 2^{\alpha_2}, \ldots\}$, and the action of $\mathfrak{S}_n$ on $\mathfrak{S}_n/\mathfrak{S}_\alpha$ corresponds to the action of $\mathfrak{S}_n$ on M_α by permuting coordinates. Hence the action of $\mathfrak{S}_n$ on $T \times T$ is equivalent to the action of $\mathfrak{S}_n$ by column permutations on the set of $2 \times n$ matrices

$$B = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{bmatrix},$$

where each row is a permutation of M_α. Associate with B the matrix A whose (i, j) entry is the number of columns of B equal to $\substack{i \\ j}$. This establishes the desired bijection between the orbits of $\mathfrak{S}_n$ acting on $T \times T$ and the $\mathbb{N}$-matrices A with $\operatorname{row}(A) = \operatorname{col}(A) = \alpha$.

7.77. a. We have

$$\left\langle \operatorname{ind}_H^G 1_H, \operatorname{ind}_K^G 1_K \right\rangle = \left\langle \operatorname{ind}_H^G 1_H \otimes \operatorname{ind}_K^G 1_K, 1_G \right\rangle.$$

Now the representation $\operatorname{ind}_H^G 1_H \otimes \operatorname{ind}_K^G 1_K$ is a *permutation* representation, obtained by letting G act on pairs (aH, bK) of left cosets by $w \cdot (aH, bK) =$

(waH, wbK). Hence the multiplicity of the trivial representation is the number of orbits of this action. When are two elements in the same orbit? Every orbit contains an element of the form (H, bK), so we are asking for what $u, v \in G$ can we find $w \in G$ such that $w \cdot (H, uK) = (H, vK)$. Since $wH = H$ we have $w \in H$, so we want to know when we can find $h \in H$ with $huk = vk'$ for $k, k' \in K$. But clearly this conditions holds if and only if $HuK = HvK$, so the number of orbits is the number of double cosets as desired.

This result can be considerably strengthened, as part of a theory developed by G. Mackey. See for example C. W. Curtis and I. Reiner, *Representation Theory of Finite Groups and Associative Algebras*, Wiley (Interscience), New York, 1962, reprinted 1988 (§10C), and [**142**, §7.3].

b. The argument given in (a) shows that the number of double cosets of (H, H) is the number of orbits of G acting diagonally on $G/H \times G/H$. This latter number is just what is meant by the rank of G acting on G/H.

c. Since $\mathrm{ch}(\mathrm{ind}_{\mathfrak{S}_\gamma}^{\mathfrak{S}_n} 1_{\mathfrak{S}_\gamma}) = h_\gamma$ (by Corollary 7.18.3), the number of double cosets of (H, K) is by (a) and Proposition 7.18.1 given by $\langle h_\alpha, h_\beta \rangle$, the number of $\mathbb{N}$-matrices A with $\mathrm{row}(A) = \alpha$ and $\mathrm{col}(A) = \beta$ (by (7.31)). The solution to Exercise 7.76(c) generalizes straightforwardly to give a combinatorial proof of the present exercise.

7.78. a. Because all characters of $\mathfrak{S}_n$ are real, we have

$$\begin{aligned} g_{\lambda\mu\nu} &= \langle \chi^\lambda \chi^\mu, \chi^\nu \rangle \\ &= \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} \chi^\lambda(w) \chi^\mu(w) \chi^\nu(w), \end{aligned}$$

which clearly has the desired symmetry.

b. We have

$$\begin{aligned} s_\lambda * s_\mu &= \sum_\nu g_{\lambda\mu\nu} s_\nu \\ &= \mathrm{ch} \sum_\nu g_{\lambda\mu\nu} \chi^\nu \\ &= \mathrm{ch}\, \chi^\lambda \chi^\mu. \end{aligned}$$

c. By linearity, we may take $f = s_\lambda$. Using (b) and the fact that $e_n = s_{1^n}$, we get

$$\begin{aligned} e_n * s_\lambda &= \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} \chi^{1^n}(w)\, \chi^\lambda(w) p_{\rho(w)} \\ &= \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} \varepsilon_w \chi^\lambda(w) p_{\rho(w)} \\ &= \omega \frac{1}{n!} \sum_{w \in \mathfrak{S}_n} \chi^\lambda(w) p_{\rho(w)} \\ &= \omega\, s_\lambda. \end{aligned}$$

d. Let $g, h \in \Lambda^n$. We claim that

$$\langle g * h, p_\nu \rangle = \langle g, p_\nu \rangle \cdot \langle h, p_\nu \rangle. \tag{7.220}$$

By bilinearity, it suffices to take $g = s_\lambda$ and $h = s_\mu$. Equation (7.220) then follows from (7.78) and (b). Now let $g = p_\lambda$ and $h = p_\mu$.

e. Note that $p_\nu(xy) = p_\nu(x)p_\nu(y)$. Hence equation (7.220) is equivalent to the desired result when $f = p_\nu$, so the general case follows by linearity.

f. One way to set up this computation is as follows:

$$\begin{aligned}
\sum_{\lambda,\mu,\nu} g_{\lambda\mu\nu} s_\lambda(x)s_\mu(y)s_\nu(z) &= \sum_{\lambda,\mu,\nu} \langle \chi^\lambda \chi^\mu, \chi^\nu \rangle s_\lambda(x)s_\mu(y)s_\nu(z) \\
&= \sum_{n\geq 0} \sum_{\lambda,\mu,\nu\vdash n} \left(\sum_{\rho\vdash n} z_\rho^{-1} \chi^\lambda(\rho)\chi^\mu(\rho)\chi^\nu(\rho) \right) \\
&\quad \times s_\lambda(x)s_\mu(y)s_\nu(z) \\
&= \sum_{n\geq 0} \sum_{\rho\vdash n} z_\rho^{-1} \left(\sum_\lambda \chi^\lambda(\rho) s_\lambda(x) \right) \\
&\quad \times \left(\sum_\mu \chi^\mu(\rho) s_\mu(x) \right) \left(\sum_\nu \chi^\nu(\rho) s_\nu(x) \right) \\
&= \sum_\rho z_\rho^{-1} p_\rho(x)p_\rho(y)p_\rho(z) \\
&= \prod_{i,j,k} (1 - x_i y_j z_k)^{-1},
\end{aligned}$$

by arguing as in the proof of Proposition 7.7.4.

g. Straightforward generalization of (f).

The internal product of symmetric functions was first defined by J. H. Redfield [**124**] (denoted something like 𝒱), and later independently by D. E. Littlewood, *J. London Math. Soc.* **31** (1956), 89–93.

7.79. a. Let xy have the meaning of Exercise 7.78(e). By the Cauchy identity (Theorem 7.12.1) applied to the two sets of variables $x_i y_j$ and z_k, we have

$$\prod_{i,j,k} (1 - x_i y_j z_k)^{-1} = \sum_\lambda s_\lambda(xy) s_\lambda(z).$$

Comparing with Exercise 7.78(f), we get

$$\begin{aligned}
s_\lambda(xy) &= \sum_\nu s_\lambda * s_\nu(x) s_\nu(y) \\
&= \sum_{\mu,\nu} g_{\lambda\mu\nu} s_\mu(x) s_\nu(y).
\end{aligned} \tag{7.221}$$

Suppose that $\langle s_\lambda, s_\mu * s_\nu \rangle = g_{\lambda\mu\nu} \neq 0$. Let $a = \ell(\mu)$, $b = \ell(\nu)$, and restrict the variables to $x = (x_1, \ldots, x_a)$ and $y = (y_1, \ldots, y_b)$. Then $s_\mu(x) \neq 0$ and $s_\nu(y) \neq 0$, so $s_\lambda(xy) \neq 0$. But $s_\lambda(xy)$ is a Schur function in the ab variables $x_i y_j$, so if $s_\lambda(xy) \neq 0$ then $\ell(\lambda) \leq ab$.

b. We can reverse the argument in (a). In equation (7.221) take $x = (x_1, \ldots, x_a)$ and $y = (y_1, \ldots, y_b)$. Since $\ell(\lambda) \leq ab$, we have $s_\lambda(xy) \neq 0$. Hence some term $g_{\lambda\mu\nu}s_\mu(x)s_\nu(y) \neq 0$, so $\ell(\mu) \leq a$ and $\ell(\nu) \leq b$. The results of this exercise (parts (a) and (b)) were first obtained (in another way) by A. Regev, *J. Algebra* **154** (1993), 125–140.

c. See Y. Dvir, *J. Algebra* **154** (1993), 125–140. For a continuation, see Y. Dvir, *Europ. J. Combinatorics* **15** (1994), 449–457.

7.80. These results are due to C. Bessenrodt and A. S. Kleshchev, On Kronecker products of complex representations of the symmetric and alternating groups, *Pacific J. Math.*, to appear.

7.81. Since $\chi^{n-1,1}(\lambda) = m_1(\lambda) - 1$ [why?], there follows for $f \in \Lambda^n$ the formula

$$f * s_{n-1,1} = p_1 \frac{\partial}{\partial p_1} f - f,$$

where we are regarding f as a polynomial in the power sums. Since

$$\frac{\partial}{\partial p_1} s_\lambda = s_{\lambda/1}$$

(see the solution to Exercise 7.35), the result follows.

7.82. a. Immediate consequence of Exercises 7.71(a)(ii) and 7.71(c).

b. This result is implicit in C. Procesi, *Advances in Math.* **19** (1976), 306–381, and *J. Algebra* **87** (1984), 342–359 (§2). An explicit statement and proof appears in A. Regev, *Linear and Multilinear Algebra* **21** (1987), 1–28 (Cor. 2.14), and 29–39 (Thm. 1).

7.83. a. We have

$$\langle \chi\psi, \phi \rangle = \frac{1}{\#G} \sum_{w \in G} \chi(w)\psi(w)\bar{\phi}(w) = \langle \chi\bar{\phi}, \bar{\psi} \rangle.$$

Hence by the Cauchy–Schwarz inequality,

$$\begin{aligned}
\langle \chi\bar{\phi}, \bar{\psi} \rangle^2 &\leq \langle \chi\bar{\phi}, \chi\bar{\phi} \rangle \, \langle \bar{\psi}, \bar{\psi} \rangle \\
&= \frac{1}{\#G} \sum_{w \in G} |\chi\phi(w)|^2 \qquad \text{(since } \psi \text{ is irreducible)} \\
&\leq \frac{1}{\#G} \phi(1)^2 \sum_{w \in G} |\chi(w)|^2 \qquad \text{(since } |\phi(w)| \leq \phi(1)\text{)} \\
&= \phi(1)^2 \langle \chi, \chi \rangle \\
&= \phi(1)^2 \qquad \text{(since } \chi \text{ is irreducible).}
\end{aligned}$$

This result appears in I. M. Isaacs, *Character Theory of Finite Groups*, Academic Press, New York, 1976; reprinted by Dover, New York, 1994 (Problem 4.12).

b. Immediate from (a), equation (7.79), and the fact that

$$\langle s_\lambda * s_\mu, s_\nu \rangle = \langle \chi^\lambda \chi^\mu, \chi^\nu \rangle.$$

7.84. a. By Exercise 7.78(e) we have

$$\begin{aligned}\langle h_\lambda * s_\mu, s_\nu\rangle &= \langle h_\lambda * s_\nu, s_\mu\rangle \qquad \text{[why?]}\\ &= \langle h_\lambda(x)s_\nu(y), s_\mu(xy)\rangle.\end{aligned}$$

Ordering the variables xy as $x_1y_1 < x_1y_2 < \cdots < x_2y_1 < x_2y_2 < \cdots$, it follows from equation (7.66) that

$$\begin{aligned}s_\mu(xy) &= \sum_{\emptyset=\mu^0\subset\mu^1\subset\cdots\subset\mu^\ell=\mu} \left(\prod_{i\geq 1} x_i s_{\mu^i/\mu^{i-1}}(y)\right)\\ &= \sum_\lambda m_\lambda(x) \sum_{\substack{\emptyset=\mu^0\subset\mu^1\subset\cdots\subset\mu^\ell=\mu\\ |\mu^i/\mu^{i-1}|=\lambda_i}} \prod_{i\geq 1} s_{\mu^i/\mu^{i-1}}(y),\end{aligned}$$

and the proof follows easily. See [**96**, Exam. I.7.23(d), p. 130].

b. Let $\nu \vdash n$. By Exercise 7.78(e) we have

$$\langle m_\nu, h_\lambda * h_\mu\rangle = \langle m_\nu(xy), h_\lambda(x)h_\mu(y)\rangle.$$

Now

$$\begin{aligned}m_\nu(xy) &= \sum_A \prod_{i,j} (x_iy_j)^{a_{ij}}\\ &= \sum_A m_{\mathrm{row}(A)}(x)m_{\mathrm{col}(A)}(y),\end{aligned}$$

where A ranges over all $\mathbb{N}$-matrices (a_{ij}) such that the decreasing rearrangement of the a_{ij}'s is ν. The proof now follows easily for the duality between the bases $\{m_\lambda\}$ and $\{h_\lambda\}$ (equation (7.30)). See [**96**, Exam. I.7.23(e), p. 131].

* **7.85.** See [**46**, Corollary 15]. For some further evaluations of $g_{\lambda\mu\nu}$, see J. B. Remmel and T. Whitehead, *Bull. Belgian Math. Soc.* **1** (1994), 649–683; E. Vallejo, On the Kronecker product of irreducible characters of the symmetric group, preprint; and the references given there. One of the main open problems in the combinatorial representation theory of $\mathfrak{S}_n$ is to obtain a combinatorial interpretation of $g_{\lambda\mu\nu}$ in general.

7.86. a. By Exercise 7.78(b) we have

$$G_{\lambda\mu}(q) = \frac{1}{n!}\sum_{w\in\mathfrak{S}_n} \frac{\chi^\lambda(w)\chi^\mu(w)}{(1-q^{\rho_1})\cdots(1-q^{\rho_\ell})}, \tag{7.222}$$

where $\rho(w) = (\rho_1, \ldots, \rho_\ell)$ with $\rho_\ell > 0$. It follows from Exercise 7.60(b) that the denominators of the nonzero terms in the above sum are all divisors of $H_\lambda(q)$, and the proof follows. This result was first given in R. Stanley, *Linear and Multilinear Algebra* **16** (1984), 29–34 (Proposition 8.2(i)). For algebraic and geometric aspects of this exercise, see P. J. Hanlon, *Adv. in Math.* **56** (1985), 238–282; J. R. Stembridge, *J. Combin. Theory (A)* **46** (1987), 79–120; R. K. Brylinski, in Lecture Notes in Math. **1404**, Springer-Verlag, Berlin/New York, 1989, pp. 35–94; and R. K. Brylinski, *Advances in Math.* **100** (1993), 28–52.

b. Multiply (7.222) by $H_\lambda(q)$ and set $q = 1$. All terms vanish except the term indexed by $w = \mathrm{id}$, yielding

$$P_{\lambda\mu}(1) = \frac{1}{n!}\chi^\lambda(\mathrm{id})\chi^\mu(\mathrm{id})H_\lambda(1) = f^\mu.$$

c. Let xy have the meaning of Exercise 7.78(e). Let $\psi = \psi_t$ be the specialization of Exercise 7.43, acting on the y variables only; and let ps_n and ps be the principal specialization and stable principal specialization of Section 7.8, acting on the x variables only. By computing directly the case $f = p_k$, we see that for any $f \in \Lambda$ and $n \in \mathbb{P}$,

$$\mathrm{ps}\,\psi_{-q^N} f(xy) = \mathrm{ps}_N\, f(x). \tag{7.223}$$

Now by Exercise 7.78(f) we have

$$s_\lambda(xy) = \sum_\mu (s_\lambda * s_\mu)(x)s_\mu(y).$$

Apply the specialization ps_N. By equation (7.223) we get

$$s_\lambda(1, q, \ldots, q^{N-1}) = \sum_\mu (s_\lambda * s_\mu)(1, q, q^2, \ldots)\,\psi_{-q^N} s_\mu(y).$$

Using Theorem 7.21.2, Corollary 7.21.3, and Exercise 7.43, we have

$$q^{b(\lambda)} \prod_{(i,j)\in\lambda} \frac{1 - q^{N+j-i}}{1 - q^{h(i,j)}} = \sum_{\mu=\langle n-k, 1^k\rangle} \frac{P_{\lambda\mu}(q)}{\prod_{(i,j)\in\lambda}\left(1 - q^{h(i,j)}\right)}(-q)^{Nk}(1-q^N).$$

Multiply by $\prod(1 - q^{h(i,j)})$ and set $t = -q^N$. We obtain a polynomial identity in t valid for infinitely many values of t (viz., $t = -q^N$) and hence valid when t is an indeterminate. Therefore

$$q^{b(\lambda)} \prod_{(i,j)\in\lambda} (1 + tq^{j-i}) = \sum_{\mu=\langle n-k, 1^k\rangle} P_{\lambda\mu}(q)t^k(1 + t),$$

which is easily seen to be equivalent to the stated result.

This result is due to A. Lascoux (private communication). His proof uses the machinery of λ-rings. Since we have not introduced this machinery here, we have given a "naive" version of Lascoux's proof. However, the λ-ring approach does give a more natural and elegant proof. For more information on λ-rings, see D. Knutson, Lecture Notes in Math. **308**, Springer-Verlag, Berlin/Heidelberg/New York, 1973.

d. This deceptively simple statement follows from [**96**, Exam. VI.8.3, pp. 362–363] and a conjecture of I. G. Macdonald, in *Actes 20ᵉ Séminaire Lotharingien*, Publ. I.R.M.A. Strasbourg, 372/S-20, 1988, pp. 131–171 (§6), and [**96**, (8.18?), p. 355]. Part (b) suggests that the coefficients of $P_{\lambda\mu}(q)$ count some statistic on SYTs of shape μ, but such an interpretation remains open. See also A. N. Kirillov, *Adv. Ser. Math. Phys.* **16** (1992), 545–579.

7.87. See Theorem 5.1 of R. Stanley, *Linear and Multilinear Algebra* **16** (1984), 3–27.

7.88. a. It follows immediately from the standard formula for an induced character (e.g., Isaacs, *ibid.* (Def. 5.1) or [**142**, Prop. 20]) that

$$\operatorname{ch} \psi_m = \frac{1}{n} \sum_{d|n} p_d^{n/d} \left(\sum_{\zeta} \zeta^m \right),$$

where ζ ranges over all primitive d-th roots of unity. The result now follows from the well-known fact (see C. A. Nicol and H. S. Vandiver, *Proc. Nat. Acad. Sci.* **40** (1954), 825–835) that the above sum over ζ is equal to

$$\frac{\phi(d)}{\phi(d/(m,d))} \mu(d/(m,d)).$$

Equation (7.189) is due to H. O. Foulkes, in *Combinatorics* (D. J. A. Welsh and D. R. Woodall, eds.), The Institute for Mathematics and Its Applications, Southend-on-Sea, Essex, 1972, pp. 141–154 (Thm. 1).

b. (sketch) Let Ω_m be the operator on $\Lambda[[q]]$ defined by

$$\Omega_m f(q) = \frac{1}{n} \sum_{\zeta^n=1} \zeta^{-m} f(\zeta q),$$

regarding n as fixed. Thus Ω_m picks out from the power series $f(q)$ those terms whose exponents are congruent to m modulo n. Apply Ω_m to the identity

$$\sum_{\lambda \vdash n} \left(\sum_{\substack{T=\text{SYT} \\ \text{of shape } \lambda}} q^{\operatorname{maj}(T)} \right) s_\lambda = \sum_{\lambda} z_\lambda^{-1} p_\lambda \frac{[n]!}{\prod [\lambda_i]},$$

where $[j] = 1 - q^j$. The coefficient of s_λ on the left-hand side is equal to the number of SYT T of shape λ satisfying $\operatorname{maj}(T) \equiv m \pmod{n}$. The coefficient of p_λ on the right-hand side is given by

$$\begin{cases} 0, & \lambda \neq \langle d^{n/d} \rangle \\ \frac{1}{n} \sum_{\substack{\zeta=\text{primitive} \\ d\text{-th root of } 1}} \zeta^m, & \lambda = \langle d^{n/d} \rangle, \end{cases}$$

and the proof follows from (a). This result was first proved independently by W. Kraśkiewicz and J. Weyman, Algebra of coinvariants and the action of Coxeter element, preprint, and by R. Stanley (unpublished). The proof by Stanley is the one given here. A similar proof appears in [**130**, Cor. 8.10]. Kraśkiewicz and Weyman extend the result to the Weyl groups of type B_n and D_n.

c. Regarding n as fixed, the expression (7.189) for ψ_m depends only on (m, n), and the proof follows from (b). A bijective proof is not known.

d. M. Kontsevich, in *The Gelfand Mathematical Seminars, 1990–1992* (L. Corwin *et al.*, eds.), Birkhäuser, Boston, 1993, pp. 173–187, mentions (p. 181) a certain representation of $\mathfrak{S}_n$ of dimension $(n-2)!$, described more explicitly (as an action on the multilinear part of the free Lie algebra on $n-1$ generators) by E. Getzler and M. M. Kapranov, in *Geometry, Topology, and Physics*, International Press, Cambridge, Massachusetts, 1995, pp. 167–201.

We will not give the definition here, but it follows from the definition that the characteristic W_n of this action is given by

$$W_n = p_1 L_{n-1} - L_n,$$

using the notation (7.191). It is easy to show from Exercise 7.89(c) that $\langle p_1 L_{n-1}, s_\lambda \rangle = y_{n-1}(\lambda)$, while Exercise 7.89(c) itself asserts that $\langle L_n, s_\lambda \rangle = y_n(\lambda)$, and the proof follows. No combinatorial proof is known.

The symmetric function W_n has subsequently appeared in a surprising number of disparate circumstances, and the $\mathfrak{S}_n$-module for which it is the characteristic is known as the *Whitehouse module*. Some references include
* E. Babson, A. Björner, S. Linusson, J. Shareshian, and V. Welker, Complexes of not i-connected graphs, MSRI preprint No. 1997-054, 31 pp.; P. Hanlon, *J. Combinatorial Theory (A)* **74** (1996), 301–320; P. Hanlon and R. Stanley, A q-deformation of a trivial symmetric group action, *Trans. Amer. Math. Soc.*, to appear; O. Mathieu, *Comm. Math. Phys.* **176** (1996), 467–474; C. A. Robinson, *Sonderforschungsbereich 343*, Universität Bielefeld, preprint 92-083, 1992; C. A. Robinson and S. Whitehouse, *J. Pure Appl. Algebra* **111** (1996), 245–253; S. Sundaram, Homotopy of non-modular partitions and the Whitehouse module, *J. Algebraic Combinatorics*, to appear; S. Sundaram, On the topology of two partition posets with forbidden block sizes, preprint, 1 May 1998; V. Turchin, Homology isomorphism of the complex of 2-connected graphs and the graph-complex of trees, preprint; S. Whitehouse, Ph.D. thesis, Warwick University, 1994; and S. Whitehouse, *J. Pure Appl. Algebra* **115** (1996), 309–321.

e. We have

$$\begin{aligned} H[J] &= \left(\sum_{i\geq 0} h_i\right)\left[\frac{\mu(d)}{d}\log(1+p_d)\right] \\ &= \left(\exp\sum_{k\geq 1}\frac{p_k}{k}\right)\left[\sum_{d\geq 1}\frac{\mu(d)}{d}\sum_{n\geq 1}(-1)^{n-1}\frac{p_d^n}{n}\right] \\ &= \exp\sum_{k\geq 1}\sum_{d\geq 1}\sum_{n\geq 1}\frac{1}{kdn}\mu(d)(-1)^{n-1}p_{kd}^n. \end{aligned}$$

Putting $N = kd$ gives

$$\begin{aligned} H[J] &= \exp\sum_{N\geq 1}\sum_{n\geq 1}\frac{1}{nN}(-1)^{n-1}p_N^n\sum_{d|N}\mu(d) \\ &= \exp\sum_{n\geq 1}(-1)^{n-1}\frac{p_1^n}{n} \\ &= \exp\log(1+p_1) \\ &= 1+p_1, \end{aligned}$$

whence $(H-1)[J] = p_1$. Since the invertible elements of $\hat{\Lambda}$ with respect to plethysm form a group, it follows also that $J[H-1] = p_1$, completing the proof.

The result of this exercise is due to C. C. Cadogan, *J. Combinatorial Theory (B)* **11** (1971), 193–200 (§3). For further aspects and a generalization, see A. R. Calderbank, P. Hanlon, and R. W. Robinson, *Proc. London Math. Soc. (3)* **53** (1986), 288–320.

7.89. a. This result can be proved by a straightforward use of the Principle of Inclusion–Exclusion. It is equivalent to enumerating primitive necklaces of length n by the number of occurrences of each color. In this form equation (7.191) appears in [**130**, Thm. 7.2].

b. This is the special case $m = 1$ of Exercise 7.88(a).

c. Let $m = 1$ in Exercise 7.88(b).

d. See e.g. [**4.21**, Thm. 5.1.5] or [**130**, (7.4.1)].

e. This result follows easily from Proposition 1.3.1.

f. This result is a consequence e.g. of I. M. Gessel and C. Reutenauer, *J. Combinatorial Theory (A)* **64** (1993), 189–215 (Thm. 3.6).

g. This key "reciprocity theorem" appears in T. Scharf and J.-Y. Thibon, *Advances in Math.* **104** (1994), 30–58 (Rmk. 3.11). A simpler proof was later given by I. M. Gessel (unpublished).

h. By (f) and (g) we have

$$\begin{aligned}\langle t_M, s_\mu\rangle &= \sum_\lambda \langle L_\lambda, h_\mu\rangle \cdot \langle p_\lambda, s_\mu\rangle \\ &= \sum_\lambda \langle p_\lambda, h_\mu\rangle \cdot \langle L_\lambda, s_\mu\rangle.\end{aligned}$$

But p_λ is clearly m-positive, so $\langle p_\lambda, h_\mu\rangle \geq 0$. Moreover, L_λ is s-positive by (b) and the fact (Appendix 2, Theorem A2.5) that the plethysm of s-positive symmetric functions remains s-positive. Hence $\langle L_\lambda, s_\mu\rangle \geq 0$, so $\langle t_M, s_\mu\rangle \geq 0$ as desired. This result was originally conjectured by R. Stanley and proved by Scharf and Thibon, *ibid.* (Example 3.15).

7.90. a. Given the SYT τ with $D(\tau) \subseteq S$, replace $1, 2, \dots, \alpha_1$ in τ by 1's; replace $\alpha_1 + 1, \alpha_1 + 2, \dots, \alpha_1 + \alpha_2$ by 2's, etc. This gives an SSYT of type $(\alpha_1, \dots, \alpha_k)$, and the correspondence is a bijection. Compare the discussion preceding Lemma 7.11.6.

b. Let τ' be the transpose of τ, of shape λ'/μ'. The condition $i \in D(\tau)$ is equivalent to $i \notin D(\tau')$, i.e., $D(\tau') \subseteq \{1, \dots, i-1, i+1, \dots, n-1\}$. Thus by (a), the number of τ of shape λ/μ with $i \in D(\tau)$ is the Kostka number $K_{\lambda'/\mu',\alpha}$, where $\alpha = (1, 1, \dots, 1, 2, 1, \dots, 1)$. But $K_{\lambda'/\mu',\alpha}$ is independent of the order of the entries of α.

7.91. a. The first statement is immediate from the Cauchy identity (Theorem 7.12.1). Similarly, if $F(t) = \prod_{i\geq 1}(1 + x_i t)$, then it follows from the dual Cauchy identity (Theorem 7.14.3) that $s_\lambda^F = s_{\lambda'}(x)$.

b. Immediate consequence of (a) and Theorem 7.21.2.

c. Write $F_{y,z}(t)$ for $F(t)$. Note that s_λ^F is a polynomial $P_\lambda(y, z)$ in y and z. When $y = 1$ and $z = q^n$ then the problem reduces to (b). Since a polynomial in one variable is determined by any infinite set of its values, it follows that (7.192) holds for $F_{1,z}(t)$. Now for any $F(t)$ let $G(t) = F(yt)$. Clearly

$s_\lambda^G = y^{|\lambda|} s_\lambda^F$. Since $F_{y,z}(t) = F_{1,z/y}(yt)$, we get

$$\begin{aligned} s_\lambda^{F_{y,z}} &= y^{|\lambda|} s_\lambda^{F_{1,z/y}} \\ &= y^{|\lambda|} q^{b(\lambda)} \prod_{u \in \lambda} \frac{1 - \frac{z}{y} q^{c(u)}}{1 - q^{h(u)}} \\ &= q^{b(\lambda)} \sum_{u \in \lambda} \frac{y - zq^{c(u)}}{1 - q^{h(u)}}. \end{aligned}$$

This result is equivalent to that of D. E. Littlewood and A. R. Richardson, *Quart. J. Math. (Oxford)* **6** (1935), 184–198 (Thm. IX) (repeated in [**88**, II. on p. 125]), and also appears in [**96**, Exam. I.3, p. 45].

d. Apply the homomorphism φ to the Jacobi–Trudi identity (Theorem 7.16.1).

e. It is trivial that (iii) ⇒ (ii), since e_λ is Schur-positive by the dual Jacobi–Trudi identity (Corollary 7.16.2).

Assume (i), so $F(t) = \prod_1^m (1 + \gamma_j t)$, where each $\gamma_j > 0$. Then

$$\prod_i F(t_i) = \prod_{j=1}^m \left(\sum_{n \geq 0} \gamma_j^n e_n(t) \right),$$

from which it follows that (i) ⇒ (iii).

Assume (i). Clearly the coefficients of $F(t)$ are then nonnegative real numbers. Let the zeros of $F(t)$ be $\theta_1, \ldots, \theta_n$, and define the Vandermonde matrix $V = (\theta_j^i)_{i,j=0}^{n-1}$. Then V is a real matrix and $A = VV^t$, so A is semidefinite. Hence (i) ⇒ (iv).

The difficult implications are (ii) ⇒ (i) and (iv) ⇒ (i). The first of these implications is equivalent to a fundamental result of M. Aissen, I. J. Schoenberg, and A. Whitney, *J. Analyse Math.* **2** (1952), 93–103. This result states that if $a_0, a_1, \ldots, a_m \in \mathbb{R}$, then every zero of the polynomial $F(t) = a_0 + a_1 t + \cdots + a_m t^m$ is a nonpositive real number if and only if every minor of the (infinite) Toeplitz matrix $A = [a_{j-i}]_{i,j \geq 0}$ (where we set $a_k = 0$ if $k < 0$ or $k > m$) is nonnegative. To see the connection with the problem under consideration, suppose that $a_0 = 1$, so that $a_i = e_i(\gamma_1, \ldots, \gamma_m)$, where $-\gamma_1^{-1}, \ldots, -\gamma_m^{-1}$ are the zeros of $F(t)$. By the dual Jacobi–Trudi identity (Corollary 7.16.2), every minor of A is a skew Schur function $s_{\lambda/\mu}(\gamma_1, \ldots, \gamma_m)$. Since skew Schur functions are s-positive (by Corollary 7.18.6 or Theorem A1.3.1), it follows that if (ii) holds then every minor of A is nonnegative. Hence by the Aissen–Schoenberg–Whitney theorem and the fact that $F(0) = 1$, every zero of $F(t)$ is a negative real number.

The above formulation of the Aissen–Schoenberg–Whitney theorem in terms of symmetric functions seems first to have been stated explicitly in R. Stanley, Graph colorings and related symmetric functions: ideas and applications, *Discrete Math.,* to appear (Thm. 2.11). An extension to arbitrary
* Toeplitz matrices (not just those with finitely many nonzero diagonals) was

given by A. Edrei, *Canad. J. Math.* **5** (1953), 86–94, and *Trans. Amer. Math. Soc.* **74** (1953), 367–383, thereby proving a conjecture of Aissen, Schoenberg, and Whitney. The same result was rediscovered by E. Thoma, *Math. Zeitschrift* **85** (1964), 40–61, in the context of the characters of the infinite symmetric group. A matrix all of whose minors are nonnegative is called a *totally positive* (or sometimes *totally nonnegative*) *matrix*. Such matrices have been extensively investigated; see, e.g., S. Karlin, *Total Positivity*, vol. 1, Stanford University Press, Stanford, California, 1968; T. Ando, *Linear Algebra Appl.* **90** (1987), 165–219; J. R. Stembridge, *Bull. London Math. Soc.* **23** (1991), 422–428; B. Kostant, *J. Amer. Math. Soc.* **8** (1995), 181–186; F. Brenti, *J. Combinatorial Theory (A)* **71** (1995), 175–218; A. D. Berenstein, S. Fomin, and A. Zelevinsky, *Advances in Math.* **122** (1996), 49–149; A. Okounkov, *Zapiski Nauchnyh Seminarov POMI* **240** (1997), 167–230; and some of the papers in *Total Positivity and Its Applications (Jaca, 1994)*, Mathematics and Its Applications **359**, Kluwer, Dordrecht, 1996. For an interesting generalization of the Aissen–Schoenberg–Whitney–Edrei–Thoma theorem, see S. Kerov, A. Okounkov, and G. Olshanski, *Internat. Math. Res. Notices* (1998), no. 4, 173–179.

The equivalence of (i)–(iii) suggests that it might be possible to prove combinatorially that certain polynomials $F(t)$ have real zeros. Assuming that $F(0) = 1$, one wants to interpret combinatorially the coefficients of the product $F(t_1)F(t_2)\cdots$ when expanded in terms of Schur functions or elementary symmetric functions, thereby showing that they are nonnegative. For an example of such an argument, see Exercise 7.47(i).

The implication (iv) $\Rightarrow$ (i) is a consequence of the work of A. Hurwitz, E. J. Routh, J. C. F. Sturm, and others on the zeros of polynomials. It seems first to have been explicitly stated by F. R. Gantmacher, *The Theory of Matrices*, vol. 2, Chelsea, New York, 1959 (Cor. on p. 203). Since a real symmetric matrix $(a_{ij})_{i,j=1}^{n}$ is semidefinite if and only if the leading principal minors $\det(a_{ij})_{i,j=1}^{k}$ are nonnegative, condition (iv) yields $n-1$ inequalities (in addition to the nonnegativity of the coefficients) on the coefficients of $F(t)$ that are necessary and sufficient for every zero of $F(t)$ to be a negative real number. (There are $n-1$ rather than n inequalities because $a_{11} = p_0^F = \deg F \geq 0$.)

7.92. **a.** See J. R. Stembridge, *Bull. London Math. Soc.* **23** (1991), 422–428. A different proof was later given by B. Kostant, *J. Amer. Math. Soc.* **8** (1995), 181–186. Note that the matrices A of this exercise are the totally positive matrices discussed in the solution to Exercise 7.91(e).

b. See J. R. Stembridge, *Canad. J. Math.* **44** (1992), 1079–1099 (Conjecture 2.1). Exercise 7.111(d) is a special case. An even stronger conjecture involving Kazhdan–Lusztig theory appears in M. Haiman, *J. Amer. Math. Soc.* **6** (1993), 569–595 (Conjecture 2.1).

7.93. Let (P, ω) be the labeled poset that is a disjoint union of chains $t_1 < \cdots < t_m$ and $t_1' < \cdots < t_n'$ with $\omega(t_i) = u_i$ and $\omega(t_j') = v_j$. It is immediate from the definition of a reverse (P, ω)-partition and from the definition (7.95) of $K_{P,\omega}$

that

$$K_{P,\omega} = L_{\mathrm{co}(u)} L_{\mathrm{co}(v)}.$$

On the other hand, we have $\mathcal{L}(P, \omega) = \mathrm{sh}(u, v)$, and the proof follows from Corollary 7.19.5.

7.94. a. Preserve the notation of Exercise 7.93. By that exercise, we have

$$\hat{\omega}(L_\alpha L_\beta) = \sum_{w \in \mathrm{sh}(u,v)} L_{\overline{\mathrm{co}(w)}}.$$

On the other hand we have

$$\begin{aligned}\hat{\omega}(L_\alpha)\hat{\omega}(L_\beta) &= L_{\bar{\alpha}} L_{\bar{\beta}} \\ &= L_{\bar{\beta}} L_{\bar{\alpha}} \\ &= \sum_{w \in \mathrm{sh}(\bar{v}, \bar{u})} L_{\mathrm{co}(w)},\end{aligned}$$

where $\bar{v} \in \mathfrak{S}_n$ and $\bar{u} \in \mathfrak{S}_{[n+1,n+m]}$ satisfy $\mathrm{co}(\bar{v}) = \bar{\beta}$ and $\mathrm{co}(\bar{u}) = \bar{\alpha}$. But then the natural bijection $\varphi : \mathrm{sh}(u, v) \to \mathrm{sh}(\bar{v}, \bar{u})$ satisfies $\overline{\mathrm{co}(w)} = \mathrm{co}(\varphi(w))$ for $w \in \mathrm{sh}(u, v)$. Hence $\hat{\omega}(L_\alpha L_\beta) = \hat{\omega}(L_\alpha)\hat{\omega}(L_\beta)$. Since the L_α's form a basis for $\mathcal{Q}$ and $\hat{\omega}$ is linear, it follows that $\hat{\omega}$ is an endomorphism of $\mathcal{Q}$. Since $\hat{\omega}$ is an involution, it is in fact an automorphism. Now let $\alpha = (1, 1, \ldots, 1) \in \mathrm{Comp}(n)$, so $\bar{\alpha} = (n)$. Then $L_\alpha = e_n$ and $L_{\bar{\alpha}} = h_n$. Hence $\hat{\omega}(e_n) = h_n$. Since $\hat{\omega}$ is an automorphism we have $\hat{\omega}(e_\lambda) = h_\lambda$ for all $\lambda \in \mathrm{Par}$, so $\hat{\omega}|_\Lambda = \omega$.

b. *First Proof.* If $C : \hat{0} = t_0 < t_1 < \cdots < t_k = \hat{1}$ is a chain of P, then write

$$f(C) = f(t_0, t_1) f(t_1, t_2) \cdots f(t_{k-1}, t_k).$$

Also for $\alpha = (\alpha_1, \ldots, \alpha_k) \in \mathrm{Comp}(n)$ let C_α denote the set of all chains $\hat{0} = t_0 < t_1 < \cdots < t_k = \hat{1}$ of P for which $\rho(t_i) - \rho(t_{i-1}) = \alpha_i$. Hence

$$F_{f^{-1}} = \sum_{\alpha \in \mathrm{Comp}(n)} \sum_{C \in C_\alpha} f^{-1}(C) M_\alpha,$$

where M_α is given by equation (7.87). Now let $f = 1 + g$, so $f^{-1} = 1 - g + g^2 - \cdots$. Then

$$\begin{aligned}f^{-1}(C) &= (1 - g + g^2 - \cdots)(t_0, t_1) \cdots (1 - g + g^2 - \cdots)(t_{k-1}, t_k) \\ &= \sum_{D \succeq C} (-1)^{\ell(D)} f(D),\end{aligned}$$

where $D \succeq C$ indicates that D is a chain refining C, and where $\ell(D)$ denotes the length of D. Hence

$$F_{f^{-1}} = \sum_{\alpha \in \mathrm{Comp}(n)} \sum_{C \in C_\alpha} \sum_{D \succeq C} (-1)^{\ell(D)} f(D) M_\alpha.$$

Now sum first over D. If $D \in C_\beta$, then α satisfies $S_\alpha \subseteq S_\beta$. Moreover,

$\ell(D) = \#S_\beta$, so we get

$$F_{f^{-1}} = \sum_{\beta \in \mathrm{Comp}(n)} \sum_{D \in C_\beta} (-1)^{\#S_\beta} f(D) \sum_{\alpha : S_\alpha \subseteq S_\beta} M_\alpha$$

$$= \sum_{\beta \in \mathrm{Comp}(n)} \sum_{D \in C_\beta} (-1)^{\#S_\beta} f(D)$$

$$\times \sum_{\alpha : S_\alpha \subseteq S_\beta} \sum_{S_\alpha \subseteq T \subseteq [n-1]} (-1)^{\#(T - S_\alpha)} L_{\mathrm{co}(T)} \qquad \text{(by (7.91))}.$$

Let $\bar{T} = [n-1] - T$. By (a) we get

$$\hat{\omega} F_{f^{-1}} = \sum_{\beta \in \mathrm{Comp}(n)} \sum_{D \in C_\beta} (-1)^{\#S_\beta} f(D) \sum_{\alpha : S_\alpha \subseteq S_\beta} \sum_{S_\alpha \subseteq T \subseteq [n-1]} (-1)^{\#(T - S_\alpha)} L_{\mathrm{co}(\bar{T})}$$

$$= \sum_{\beta \in \mathrm{Comp}(n)} \sum_{D \in C_\beta} (-1)^{\#S_\beta} f(D) \sum_{T \subseteq [n-1]} L_{\mathrm{co}(\bar{T})} \sum_{\alpha : S_\alpha \subseteq S_\beta \cap T} (-1)^{\#(T - S_\alpha)}.$$

But

$$\sum_{\alpha : S_\alpha \subseteq S_\beta \cap T} (-1)^{\#(T - S_\alpha)} = \begin{cases} (-1)^{\#T} & \text{if } S_\beta \cap T = \emptyset \\ 0 & \text{otherwise.} \end{cases}$$

Since $S_\beta \cap T = \emptyset$ if and only if $S_\beta \subseteq \bar{T}$, we get

$$\hat{\omega} F_{f^{-1}} = \sum_{\beta \in \mathrm{Comp}(n)} \sum_{D \in C_\beta} (-1)^{\#S_\beta} f(D) \sum_{S_\beta \subseteq \bar{T} \subseteq [n-1]} (-1)^{n - \#\bar{T}} L_{\mathrm{co}(\bar{T})}$$

$$= (-1)^n \sum_{\beta \in \mathrm{Comp}(n)} \sum_{D \in C_\beta} f(D) M_{\mathrm{co}(\beta)} \qquad \text{(by (7.91))}$$

$$= (-1)^n F_f,$$

completing the proof.

Second Proof (Sketch). Let $m \in \mathbb{P}$. We see immediately from the relevant definitions that

$$F_{f^m}(x) = F_f(mx), \tag{7.224}$$

where mx denotes the multiset of variables consisting of m x_1's, m x_2's, etc., in that order. If for any $G \in \mathcal{Q}^n$ we expand $g(mx)$ in terms of some basis for $\mathcal{Q}^n$, then the coefficients will be polynomials in m. One can show (using the basis $\{L_\alpha\}$) that setting $m = -1$ yields $(-1)^n \hat{\omega}(G)$. (Compare equation (A2.163) of Appendix 2.) Similarly if we expand the left-hand side of (7.224) in terms of some basis for $\mathcal{Q}^n$, then the coefficients will be polynomials in m, and setting $m = -1$ yields $F_{f^{-1}}(x)$. Hence the proof follows by setting $m = -1$ in (7.224).

7.95. a. This is a straightforward generalization of Lemma 7.23.3. Regarding P_S as the border strip B_α, we have that a permutation $v \in \mathfrak{S}_n$ belongs to

$\mathcal{L}(P_S, \omega_w)$ if and only if w_{i+1} follows w_i in v whenever w_i is to the left of w_{i+1} in the same row, and w_{i+1} precedes w_i in v whenever w_i is below w_{i+1} in the same column. This condition is easily seen to be equivalent to $D(wv^{-1}) = S$.

b. The set $\mathcal{A}^r(P_S, \omega_w)$ of reverse (P_S, ω_w)-partitions depends on S and on the set of those pairs $(u, v) \in P_S \times P_S$ for which $u < v$ and $\omega_w(u) > \omega_w(v)$. This set of pairs in turn depends only on $D(w)$. Hence by Corollary 7.19.5 the multiset $M = \{D(v) : v \in \mathcal{L}(P_S, \omega_w)\}$ depends only on S and $D(w)$. By (a), we have $M = \{D(v) : D(wv^{-1}) = S\}$. Letting $u = wv^{-1}$, we get that the number of u, v for which $D(u) = S$, $D(v) = T$, and $uv = w$ depends only on u, v, and $D(w)$, as was to be proved.

This result may be formulated algebraically as follows. In the group algebra $\mathbb{Q}\mathfrak{S}_n$ of $\mathfrak{S}_n$, define for each $S \subseteq [n-1]$

$$B_S = \sum_{w\,:\,D(w)=S} w.$$

Then the algebra $\mathcal{D}_n$ generated by the B_S's is equal (as a set) to their linear span. In other words, $\dim \mathcal{D}_n = 2^{n-1}$. The algebra $\mathcal{D}_n$ is known as the *descent algebra* of $\mathfrak{S}_n$ and has many remarkable properties. It was first defined (for any finite Coxeter group) by L. Solomon, *J. Algebra* **41** (1976), 255–268. For a connection between descent algebras and quasisymmetric functions, see C. Malvenuto and C. Reutenauer, *J. Algebra* **177** (1995), 967–982. The proof that $\dim \mathcal{D}_n = 2^{n-1}$ given here is due to I. M. Gessel. A good reference to the descent algebra is [**130**, Ch. 9].

7.96. This result is implicit in A. M. Garsia and C. Reutenauer, *Advances in Math.* **77** (1989), 189–262 (Thm. 4.4).

7.97. **a.** It follows from the four equations beginning with (7.98) that

$$F(x) = \sum_{P,Q} x^{|P|+|Q|-|\lambda|}$$
$$= x^{-t}\left(\sum_P x^{|P|}\right)\left(\sum_Q x^{|Q|}\right),$$

where P ranges over all reverse SSYT of shape λ and largest part at most c, while Q ranges over all reverse SSYT of shape λ and largest part at most r. The sum over P is thus just $s_\lambda(x, x^2, \ldots, x^r)$, while the sum over Q is $s_\lambda(x, x^2, \ldots, x^c)$, and the proof follows.

b. The proof is parallel to that of (a), using the correspondence $\pi \mapsto \sigma$ defined in the second proof of Theorem 7.20.4.

7.98. **a.** This formula can be proved by generalizing the proof of Theorem 7.22.1 (the Hillman–Grassl algorithm). See E. R. Gansner, *J. Combinatorial Theory (A)* **30** (1981), 71–89 (Thm. 5.1).

b. See the above reference (Thm. 6.1).

7.99. From the proof of Theorem 7.20.1 we have

$$\sum_{t,n\geq 0} K_t(n) q^t x^n = \sum_{a_{ij}} q^{\sum a_{ij}} x^{\sum (i+j-1)a_{ij}},$$

where each a_{ij} for $i, j \geq 1$ ranges over the set $\mathbb{N}$. Putting q/x for q gives

$$\sum_{t,n\geq 0} K_t(n+t)q^t x^n = \sum_{a_{ij}} q^{\sum a_{ij}} x^{\sum (i+j-2)a_{ij}}.$$

The condition that $n \leq t$ becomes

$$\sum (i+j-2)a_{ij} \leq \sum a_{ij},$$

or equivalently,

$$a_{11} \geq \sum{}' (i+j-3)a_{ij},$$

where $\sum'$ indicates that the term $(i, j) = (1, 1)$ is missing. Hence

$$\begin{aligned}
\sum_{0\leq n\leq t} K_t(n+t)q^t x^n &= \sum_{a_{11}\geq \sum'(i+j-3)a_{ij}} q^{\sum a_{ij}} x^{\sum (i+j-2)a_{ij}} \\
&= \frac{1}{1-q} \sum_{a_{ij}:(i,j)\neq(1,1)} q^{\sum (i+j-2)a_{ij}} x^{\sum (i+j-2)a_{ij}} \\
&= \frac{1}{1-q} \prod_{i,j}{}' \left(\sum_{a_{ij}\geq 0} (qx)^{(i+j-2)a_{ij}} \right) \\
&= \frac{1}{1-q} \prod_{i,j}{}' \frac{1}{1-(qx)^{i+j-2}} \\
&= \frac{1}{1-q} \prod_{k\geq 1} \frac{1}{[1-(qx)^k]^{k+1}}.
\end{aligned}$$

From this the desired conclusion is immediate.

This result was first proved (by a more complicated method) in R. Stanley, *J. Combinatorial Theory (A)* **14** (1973), 53–64 (Cor. 5.3(v)).

The result of this exercise is equivalent to the formula

$$K_t(n+t) = \sum_{k=0}^{n} p(k)a(n-k), \qquad 0 \leq n \leq t,$$

where $p(k)$ denotes the number of partitions of k and $a(n-k)$ the number of plane partitions of $n-k$. Is there a direct bijective proof?

7.100. **a.** Let $A \xrightarrow{\text{RSK}} (P, Q)$. Proposition 7.23.10 tells us the first row of P. Theorem 7.23.16 then allows us to describe the first column of P directly in terms of A. Using Theorem 7.13.1 we can then describe the first column of Q also in terms of A. These are all the ingredients necessary for the proof, though we omit the details.

b. The result of (a) applies equally well to the "reverse" version of the RSK algorithm used in the proof of Theorem 7.20.1. Hence the merged plane partitions $\pi(P, Q)$ and $\pi(P', Q')$ (as defined in the proof of Theorem 7.20.1) will have the same first column, so the row conjugates $\pi'(P, Q)$ and

$\pi'(P', Q')$ will have the same shape. In other words, the shape $\mathrm{sh}(\pi')$ depends only on the support $\mathrm{supp}(A)$. Since under the correspondence $A \mapsto \pi'$ we have $\mathrm{tr}(\pi') = \sum a_{ij}$, it follows that there is a collection $\mathcal{S}$ of finite subsets of $\mathbb{P} \times \mathbb{P}$ such that the following condition holds: $\mathrm{sh}(\pi') \subseteq \lambda$ and $\mathrm{tr}(\pi') = n$ if and only if $\mathrm{supp}(A) \in \mathcal{S}$ and $\sum a_{ij} = n$. If $S \subset \mathbb{P}^2$ with $\#S = k$, then the number of $\mathbb{N}$-matrices of support S and element sum n is just $\binom{n-1}{k-1}$, the number of compositions of n into k parts. Thus we get

$$t_\lambda(n) = \sum_{S \in \mathcal{S}} \binom{n-1}{\#S-1},$$

which is a polynomial in n. Given that $t_\lambda(n)$ is a polynomial, there are several ways to see that its degree is $|\lambda|-1$. For instance, Theorem 4.5.4 (or the more general Corollary 7.19.5) allows the generating function $\sum_n t_\lambda(n)x^n$ to be expressed as a sum of f^λ rational functions of the form $x^a/(1-x^{b_1})\cdots(1-x^{b_m})$ where $\lambda \vdash m$, from which it is immediate that $\deg t_\lambda = |\lambda| - 1$. See E. R. Gansner, *Illinois J. Math.* **25** (1981), 533–554.

c. The condition that π' fits in an $a \times b$ rectangle is equivalent (by the proof of Theorem 7.20.1) to $\max\{j : (i, j) \in \mathrm{supp}(A)\} \le b$ and $\max\{i : (i, j) \in \mathrm{supp}(A)\} \le a$. Hence $t_{\langle b^a \rangle}(n)$ is equal to the number of $a \times b$ $\mathbb{N}$-matrices whose entries sum to n, which is just $\binom{ab+n-1}{ab-1}$.

7.101. Equation (7.193) was stated without proof (with a misprint) by R. Stanley, in *Combinatoire et Représentation du Groupe Symétrique (Strasbourg 1976)*, Lecture Notes in Math. **579**, Springer-Verlag, Berlin/Heidelberg/New York, 1977, pp. 217–251 (Thm. 4.3(b)). The first proof was given by R. A. Proctor, in *Lie Algebras and Related Topics* (D. J. Britten, F. W. Lemire, and R. V. Moody, eds.), CMS Conf. Proc. **5**, American Mathematical Society, Providence, 1986, pp. 357–360, and *Invent. Math.* **92** (1988), 307–332 (Cor. 4.1). Proctor actually proves the case $d = 1$ of equation (7.194), and later states (immediately after Cor. 4.1) equation (7.194) in its full generality. Proctor's proof is based on representation theory; the number $f_n(m)$ is in fact the dimension of the irreducible representation of the symplectic group $\mathrm{Sp}(2(n-1), \mathbb{C})$ (or Lie algebra $\mathfrak{sp}(2(n-1), \mathbb{C})$) with highest weight $m\lambda_{n-1}$, where λ_{n-1} denotes the $(n-1)$-st fundamental weight. (See also Exercise 6.25(c).) Proctor proves the more general case $d = 1$ of (7.194) also using representation theory, but when $M - \ell$ is even this involves the construction of a non-semisimple analogue $\mathfrak{sp}(2n+1, \mathbb{C})$ of the symplectic Lie algebra $\mathfrak{sp}(2n, \mathbb{C})$. Proctor's unpublished proof of the general case of (7.194) uses entirely different techniques, viz., the evaluation of the $q = 1$ case of MacMahon's determinantal expression (P. A. MacMahon, *Phil. Trans. Roy. Soc. London (A)* **211** (1911), 345–373 (p. 367); *Collected Works*, vol. 1 (G. E. Andrews, ed.), MIT Press, Cambridge, Massachusetts, 1978, pp. 1406–1434 (p. 1428)) for the polynomial $\sum_\pi q^{|\pi|}$, summed over all plane partitions, allowing 0 as a part, of an arbitrary shape μ. Subsequently a much more general determinant evaluation was given by C. Krattenthaler, *Manuscripta Math.* **69** (1990), 173–202, and in *Number-Theoretic Analysis* (H. Hlawka and R. F. Tichy, eds.), Lecture Notes in Math. **1452**, Springer-Verlag, Berlin/Heidelberg/New York, 1990.

pp. 121–131 (though in the special case of (7.194) Krattenthaler does not state the result in the elegant form we have given, due to Proctor). For some further applications of representation theory to the enumeration of plane partitions, see R. A. Proctor, *Europ. J. Combinatorics* **11** (1990), 289–300.

7.102. a. By Theorem 7.21.2 and Corollary 7.21.3 we have

$$t_\lambda(q) = s_\lambda(1, q, \ldots, q^n),$$

which clearly has the desired properties.

b. By the Jacobi–Trudi identity (Theorem 7.16.1) and the specialization $h_i(1, q, \ldots) = 1/[i]!$ (where $[i]! = (1-q)(1-q^2)\cdots(1-q^i)$), we have

$$\begin{aligned} t_{\lambda/\mu,n}(q) &= \prod_{u\in\lambda/\mu} \left(1 - q^{n+c(u)}\right) \cdot \det\left(\frac{1}{[\lambda_i - \mu_j - i + j]!}\right) \\ &= \prod_{u\in\lambda/\mu} \left(1 - q^{n+c(u)}\right) \cdot \sum_{w\in\mathfrak{S}_n} \varepsilon_w \prod_i \frac{1}{\left[\lambda_i - \mu_{w(i)} - i + w(i)\right]!}. \end{aligned}$$

It is not hard to see that

$$\prod_{u\in\lambda/\mu} \left(1 - q^{n+c(u)}\right) \cdot \prod_i \frac{1}{[\lambda_i - \mu_j - i + j]!} = \prod_i \binom{n + \lambda_i - i}{\lambda_i - \mu_j - i + j}.$$

It follows that

$$t_{\lambda/\mu,n} = \det\left(\binom{n + \lambda_i - i}{\lambda_i - \mu_j - i + j}\right). \tag{7.225}$$

The proof now follows from a straightforward application of Theorem 2.7.1. (In fact, equation (7.225) is a specialization of a result known as the "Jacobi–Trudi identity for flag Schur functions," due to I. M. Gessel and appearing in M. L. Wachs, *J. Combinatorial Theory (A)* **40** (1985), 276–289 (Thm. 3.5).) The proof we have just given is due to H. L. Wolfgang (private communication, 13 November 1996). A version of the proof, based on the theory of Schubert polynomials, appears in S. C. Billey, W. Jockusch, and R. Stanley, *J. Algebraic Combinatorics* **2** (1993), 345–375 (Thm. 3.1). Is there a "nice" bijective proof?

7.103. a. This result was conjectured by I. G. Macdonald and proved by J. R. Stembridge, *Advances in Math.* **111** (1995), 227–243.

b. This result was conjectured by D. P. Robbins and R. Stanley, and proved by G. E. Andrews, *J. Combinatorial Theory (A)* **66** (1994), 28–39.

Both (a) and (b) (as well as Theorem 7.20.4 and Exercise 7.106(b)) are part of the subject of the enumeration of symmetry classes of plane partitions. For an overview of this subject (written when most of the current theorems were conjectures), see R. Stanley, *J. Combinatorial Theory (A)* **43** (1986), 103–113; Erratum, **44** (1987), 310. All ten symmetry classes discussed in this paper have now been enumerated, though the q-enumeration of totally symmetric plane partitions (i.e., the plane partitions of (a)) remains open. For a recent paper with further references, see G. Kuperberg, *J. Combinatorial*

Theory (A) **75** (1996), 295–315. For an entertaining account of the numbers $B(r)$ of (b), see D. P. Robbins, *Math. Intelligencer* **13** (1991), 12–19.

c. This intriguing and surprisingly difficult result was conjectured by W. H. Mills, D. P. Robbins, and H. Rumsey, Jr., *Invent. Math.* **66** (1982), 73–87 (Conjecture 1), and *J. Combinatorial Theory (A)* **34** (1983), 340–359 (Conjecture 1). It was first proved by D. Zeilberger, *Electron. J. Combinatorics* **3** (1996), R13, 84 pp.; also published in *The Foata Festschrift* (J. Désarménien, A. Kerber, and V. Strehl, eds.), Imprimerie Louis-Jean, Gap, France, 1996, pp. 289–372. A simpler proof was later given by G. Kuperberg, *Int. Math. Res. Notices (3)*, 1996, 139–150. For more information concerning this result, see the paper of Robbins cited in (b). For a textbook devoted to symmetry classes of plane partitions, monotone triangles, and related topics, see D. M. Bressoud, *Proofs and Confirmations: The Story of the Alternating Sign Matrix Conjecture*, Cambridge University Press and Mathematical Association of America, to appear.

d. If $w = w_1 \cdots w_n \in \mathfrak{S}_n$, then associate with w the monotone triangle $\mathrm{mt}(w) = (a_{ij}(w))_{1\leq i\leq j\leq n}$ whose i-th row consists of the numbers $w_1, w_2, \ldots, w_{n-i+1}$ in increasing order. It was shown by C. Ehresmann, *Ann. Math.* **35** (1934), 396–443 (§20) that the set $\{\mathrm{mt}(w) : w \in \mathfrak{S}_n\}$, ordered componentwise, is isomorphic to P_n. Given triangular arrays $a = (a_{ij})$ and $b = (b_{ij})$, define the *meet* $a \wedge b$ to be the triangular array $(\min\{a_{ij}, b_{ij}\})$. It is not difficult to check that the set of all arrays obtained by repeatedly taking meets of the triangles $\mathrm{mt}(w)$, $w \in \mathfrak{S}_n$, coincides with the set of monotone triangles. Hence $L(P_n)$ is a completion of P_n, and it is not hard to show that it is in fact the MacNeille completion. The surprising formula $\#L(P_n) = M(n)$ is due to A. Lascoux and M. P. Schützenberger, *Electron. J. Combinatorics* **3** (1996), R27, 35 pp.; also published in *The Foata Festschrift* (J. Désarménien, A. Kerber, and V. Strehl, eds.), Imprimerie Louis-Jean, Gap, France, 1996, pp. 653–685. Note also the unexpected fact that $L(P_n)$ is a distributive lattice.

7.104. This is a result of E. M. Wright, *Quart. J. Math. Oxford (2)* **2** (1931), 177–189.

7.105. Yes if $n < 14$, but no for $n = 14$, an example being $\lambda = (5, 5, 2, 1, 1)$ and $\mu = (4, 4, 3, 1, 1, 1)$. These results are due to L. A. Shepp, private communication, 1975.

7.106. a. It is a straightforward application of the Littlewood–Richardson rule (Appendix 1, Section A1.3) that

$$s_\nu^2 = \sum_{\lambda\subseteq\langle c^r\rangle} s_{(c+\lambda_1, c+\lambda_2, \ldots, c+\lambda_r, c-\lambda_r, c-\lambda_{r-1}, \ldots, c-\lambda_1)}. \tag{7.226}$$

A direct bijective proof using jeu de taquin (essentially a proof of the Littlewood–Richardson rule in the special case s_ν^2) can also be given; see R. Stanley, *J. Combinatorial Theory (A)* **43** (1986), 103–113; Erratum, **44** (1987), 310.

b. Let π be a $(2r, 2c, 2t)$-self-complementary plane partition. Add $2r - i$ to every entry of the i-th row of π to obtain a reverse column-strict plane partition σ of shape $\langle(2c)^{2r}\rangle$, allowing 0 as a part, with largest part at most $2r + 2t - 1$.

Let τ be the subarray of σ consisting of all entries less than $r+t$. Then τ is a (rotated) SSYT (allowing 0 as a part) with largest part at most $r+t-1$ and whose shape is one of the partitions μ such that s_μ is a term of the sum in equation (7.226). Conversely, given such an SSYT, we can reverse the steps to obtain a $(2r, 2c, 2t)$-self-complementary plane partition. It follows that

$$G(2r, 2c, 2t) = s_{\langle c^r\rangle}(1^{r+t})^2.$$

But $s_{\langle c^r\rangle}(1^{r+t}) = F(r, c, t)$, since an SSYT with $\leq r$ rows, $\leq c$ columns, and largest part $\leq r+t-1$ (allowing 0 as a part) can be converted to a plane partition with $\leq r$ rows, $\leq c$ columns, and largest part $\leq r+t$ by subtracting $r-i$ from the entries in the i-th row and rotating 180°. Hence we get equation (7.195).

In a similar manner we obtain

$$G(2r+1, 2c, 2t) = F(r, c, t)F(r+1, c, t)$$
$$G(2r+1, 2c+1, t) = F(r+1, c, t)F(r, c+1, t).$$

These results appear in R. Stanley, *ibid.* (eqns. (3a)–(3c)).

7.107. a. Let $\mu^{[n]}$ denote the partition whose diagram is an $n \times n$ square (with n sufficiently large) with the shape μ removed from the bottom right-hand corner. By Exercise 7.41 we have

$$s_{\mu^{[n]}}(x_1, \ldots, x_n) = (x_1 \cdots x_n)^n s_\mu\big(x_1^{-1}, \ldots, x_n^{-1}\big).$$

Put $x_i = q^{i-1}$. By Theorem 7.21.2 we get

$$q^k \prod_{u \in \mu^{[n]}} \frac{1-q^{n+c(u)}}{1-q^{h(u)}} = \prod_{v \in \mu} \frac{1-q^{-n-c(v)}}{1-q^{-h(v)}}$$

for some $k \in \mathbb{Z}$. The right-hand side is equal to $q^m \prod_{v \in \mu}(1-q^{n+c(v)})/(1-q^{h(v)})$ for some $m \in \mathbb{Z}$, so

$$q^{k-m} \prod_{u \in \mu^{[n]}} \frac{1-q^{n+c(u)}}{1-q^{h(u)}} = \prod_{v \in \mu} \frac{1-q^{n+c(v)}}{1-q^{h(v)}}.$$

Putting $q=0$ shows that $k-m=0$. Now as $n \to \infty$, it is easily seen that

$$\prod_{u \in \mu^{[n]}} \big(1-q^{n+c(u)}\big) \to \prod_{i \geq 1}(1-q^i)^i$$
$$\prod_{u \in \mu^{[n]}} \big(1-q^{h(u)}\big) \to \prod_{u \in A_\mu} \big(1-q^{h(u)}\big)$$
$$\prod_{v \in \mu} \big(1-q^{n+c(v)}\big) \to 1,$$

and the proof follows.

A bijective proof of this exercise was first given by D. E. White (private communication). We sketch another such proof (found in collaboration

with C. Bessenrodt) using the binary sequence coding C_μ of μ explained in Exercise 7.59. By considering the sequences C_μ and C_{A_μ} together with Exercise 7.59(a), one sees that we need to prove bijectively the following.

Lemma. *Let C be a binary sequence $\cdots c_{-1}c_0c_1\cdots$ beginning with infinitely many 0's and ending with infinitely many 1's. For each $p \geq 1$, let $r_p(C)$ be the number of integers i such that $c_i = 0$ and $c_{i+p} = 1$, and let $s_p(C)$ be the number of integers i such that $c_i=1$ and $c_{i+p} = 0$. Then $r_p(C) = p + s_p(C)$.*

Proof. First note that the case $p = 1$ is easy to prove bijectively. But when we apply the case $p = 1$ to each of the subsequences $C^j = \{c_{pi+j}\}_{i\in\mathbb{Z}}$, where $0 \leq j < p$, then we obtain the stated result. □

C. Bessenrodt has observed that the present exercise is equivalent to the statement that for each $p \geq 1$, the number of ways to add a border strip of size p to μ is exactly p more than the number of border strips of μ of size p. Note that the case $p = 1$ is the familiar fact (see Exercise 3.22(c)) that in Young's lattice the element μ is covered by one more element than it covers. The general case then follows from the isomorphism $Y_{p,\emptyset} \xrightarrow{\cong} Y^p$
* of Exercise 7.59(e). For related work see C. Bessenrodt, On hooks of Young diagrams, preprint.

b. A weak *reverse* plane partition of shape $\mu^{[n]}$, rotated 180°, is just a skew plane partition of shape $\langle n^n\rangle/\mu$. Hence by Theorem 7.22.1,

$$\sum_\pi q^{|\pi|} = \frac{1}{\prod_{u\in\langle n^n\rangle/\mu}[h(u)]}.$$

Now let $n \to \infty$ and use (a).

c. Such a proof was given by K. Kadell, *J. Combinatorial Theory (A)* **77** (1997), 110–133 (§6).

7.108. The only partition of $p+q$ with largest part p and with q parts is $\langle p, 2, 1^{q-2}\rangle$. Hence by Theorem 7.23.13, we have

$$\begin{aligned} F(p,q) &= \left(f^{\langle p,2,1^{q-2}\rangle}\right)^2 \\ &= \left(\frac{(p+q)!}{pq(p+q-1)(p-2)!(q-2)!}\right)^2. \end{aligned}$$

7.109. a. Immediate from Corollary 7.23.12.

b. This result was first shown by J. M. Hammersley, in *Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, University of California Press, Berkeley/Los Angeles, 1972, pp. 345–394 (Thm. 4), using subadditive ergodic theory.

c. If $w = w_1w_2\cdots w_n \in \mathfrak{S}_n$, then let $w^r = w_n\cdots w_2w_1$. It follows from Example 7.23.19(a) that $\mathrm{is}(w)\cdot\mathrm{is}(w^r) \geq n$. Hence (using the arithmetic–

geometric-mean inequality),

$$\begin{aligned} E(n) &= \frac{1}{n!}\frac{1}{2}\sum_{w\in\mathfrak{S}_n}[\mathrm{is}(w)+\mathrm{is}(w^r)] \\ &\geq \frac{1}{n!}\sum_{w\in\mathfrak{S}_n}\sqrt{\mathrm{is}(w)\cdot\mathrm{is}(w^r)} \\ &\geq \frac{1}{n!}\sum_{w\in\mathfrak{S}_n}\sqrt{n} \\ &= \sqrt{n}, \end{aligned}$$

and the proof follows. By a more sophisticated argument Hammersley, *ibid.* (p. 360), shows that $\alpha \geq \pi/2 = 1.57\cdots$, and explains how this bound can be improved to $\sqrt{8/\pi} = 1.59\cdots$ (also done independently by D. H. Blackwell).

d. The number of subsequences of $w \in \mathfrak{S}_n$ of length k is $\binom{n}{k}$, and the probability that a specified one of them is increasing is $1/k!$. Hence the probability that $\mathrm{is}(w) \geq k$ cannot exceed $\frac{1}{k!}\binom{n}{k}$, and the proof can be completed by a judicious use of Stirling's formula. This argument is due to Hammersley, *ibid.* (Thm. 6).

e. This result was proved independently by B. F. Logan and L. A. Shepp, *Advances in Math.* **26** (1977), 206–222, and by A. M. Vershik and S. V. Kerov, *Dokl. Akad. Nauk SSSR* **233** (1977), 1024–1027, English translation in *Soviet Math. Dokl.* **18** (1977), 527–531.

f. Roughly speaking, most of the contribution to the sum on the right-hand side of (7.196) comes from terms indexed by λ "near" $\tilde{\lambda}^n$. Moreover, since $\sum_{\lambda\vdash n}(f^\lambda)^2 = n!$ and the number of terms of this sum is small compared with $n!$, we see that $(f^{\tilde{\lambda}^n})^2$ is "near" $n!$. Thus the largest part $(\tilde{\lambda}^n)_1$ of $\tilde{\lambda}^n$ is "near" $\alpha\sqrt{n}$. Since $\lim_{x\to 0} f(x) = 2$, it follows that $(\tilde{\lambda}^n)_1$ is asymptotically at least as large as $2\sqrt{n}$, so $\alpha \geq 2$. For rigorous treatments of this argument, see the two papers cited in (e) above.

g. See A. M. Vershik and S. V. Kerov, *ibid.*

Much further work has been subsequently done on the problems of estimating $E(n)$ and describing $\tilde{\lambda}^n$, and the closely related problem of finding the "typical" shape of a permutation $w \in \mathfrak{S}_n$ (i.e., the shape of the SYT P or Q obtained from w *via* the RSK algorithm). See for instance S. V. Kerov and A. M. Vershik, *SIAM J. Alg. Disc. Meth.* **7** (1986), 116–124; J. M. Steele, in *Discrete Probability and Applications (Minneapolis, MN, 1993)*, IMA Vol. Math. Appl. **72**, Springer, New York, 1995, pp. 111–131; D. Aldous and P. Diaconis, *Probab. Theory Related Fields* **103** (1995), 199–213; J. H. Kim, *J. Combinatorial Theory (A)* **76** (1996), 148–155; T. Seppäläinen, *Electron. J. Probab.* **1** (1996), no. 5, 51 pp.; B. Bollobás and S. Janson, in *Combinatorics, Geometry and Probability (Cambridge, 1993)*, Cambridge University Press, Cambridge, 1997, pp. 121–128; and J. Baik, P. Deift, and K. Johansson, On the distribution of the length of the longest increasing subsequence of a random permutation, preprint dated May 11, 1998. In this last paper the

following remarkable result is proved. Let $u(x)$ be the (unique) solution of the Painlevé II equation

$$u_{xx} = 2u^3 + xu, \quad \text{and } u \sim \mathrm{Ai}(x) \text{ as } x \to \infty,$$

where $\mathrm{Ai}(x)$ is the Airy function. Let L_n denote the length of the longest increasing subsequence of a random permutation $w \in \mathfrak{S}_n$. Then

$$\lim_{n\to\infty} \mathrm{Prob}\left(\frac{L_n - 2\sqrt{n}}{n^{1/6}} \le t\right) = \exp\left(-\int_t^\infty (x-t)u(x)^2 dx\right).$$

In particular, the variance of L_n is an explicit constant times $n^{1/3}$.

7.110. From equation (7.96) there follows

$$\begin{aligned}
\frac{1-q}{q}\sum_{m\ge 0} s_\lambda(\underbrace{1-q,\dots,1-q}_{m})q^m
&= \frac{\sum_{\mathrm{sh}\,T=\lambda} q^{d(T)+1}}{(1-q)^{|\lambda|+1}} \\
\Longrightarrow\quad Z = \sum_\lambda \left(\sum_{\mathrm{sh}\,T=\lambda} q^{d(T)}\right) s_\lambda
&= \frac{1-q}{q}\sum_\lambda s_\lambda \sum_{m\ge 0} s_\lambda(1-q,\dots,1-q)q^m \\
&= \frac{1-q}{q}\sum_{m\ge 0} q^m \sum_\lambda s_\lambda \cdot s_\lambda(1-q,\dots,1-q) \\
&= \frac{1-q}{q}\sum_{m\ge 0} q^m \prod_i \frac{1}{[1-x_i(1-q)]^m},
\end{aligned}$$

the last step by the Cauchy identity (Theorem 7.12.1). Note that

$$\prod_i \frac{1}{1-x_i(1-q)} = \sum_{n\ge 0}(1-q)^n s_n.$$

Hence

$$Z = \frac{1-q}{q}\,\frac{1}{1-\frac{1}{\sum_{n\ge 0}(1-q)^n s_n}} = \frac{\sum_{n\ge 0}(1-q)^n s_n}{1-q\sum_{n\ge 1}(1-q)^{n-1}s_n}. \tag{7.227}$$

Now note that from the generating function for the Eulerian polynomials given in Exercise 3.81(c) we get

$$\begin{aligned}
1 + \frac{1}{q}\sum_{\ell\ge 1} A_\ell(q)\frac{x^\ell}{\ell!} &= \frac{(1-q)e^{(1-q)x}}{1-qe^{(1-q)x}} \\
&= \frac{e^{(1-q)x}}{1-q\sum_{\ell\ge 1}(1-q)^{\ell-1}\frac{x^\ell}{\ell!}}.
\end{aligned}$$

Letting $H(z) = \sum_{n\geq 0} s_n z^n$, the right-hand side of (7.227) can be written as

$$Z = \frac{(1-q)H(1-q)}{1-qH(1-q)}.$$

By equation (7.198) we have

$$H(z) = \exp\left(p_1 z + p_2 \frac{z^2}{2} + \cdots\right),$$

so we get

$$Z = \frac{1}{q} \sum_{\ell \geq 1} A_\ell(q) \frac{\left[p_1(1-q) + \frac{1}{2}p_2(1-q)^2 + \cdots\right]^\ell}{\ell!}.$$

Expanding by the multinomial theorem shows that the coefficient of p_λ is

$$z_\lambda^{-1} q^{-1} (1-q)^{n-\ell} A_\ell(q) p_\lambda,$$

as desired. This argument was done in collaboration with I. M. Gessel.

7.111. **a.** We have $X = \mathfrak{S}_n$. Hence the formula $Z_X = n! h_n$ is equivalent to equation (7.22).

b. Let $Y_i = \{w \in \mathfrak{S}_n : w(n) = i\}$. Clearly $\tilde{Z}_{Y_1} = \tilde{Z}_{Y_2} = \cdots = \tilde{Z}_{Y_{n-1}}$, while by (a) we have $\tilde{Z}_{Y_n} = (n-1)!\, h_{n-1} h_1$. But

$$n!\, h_n = \tilde{Z}_{[n]\times[n]} = \tilde{Z}_{Y_1} + \cdots + \tilde{Z}_{Y_n},$$

so

$$\begin{aligned}
\tilde{Z}_X &= \tilde{Z}_{[n]\times[n]} - \tilde{Z}_{Y_1} \\
&= n!\, h_n - \frac{1}{n-1}[n!\, h_n - (n-1)!\, h_{n-1} h_1] \\
&= n(n-2)(n-2)!\, h_n + (n-2)!\, h_{n-1} h_1.
\end{aligned}$$

c. The h-positivity of $\tilde{Z}_X$ is equivalent to Exercise 7.47(l) in the case when the complement of G is bipartite. See the solution to that exercise for references. It follows from this solution that moreover the only h_λ's appearing in $\tilde{Z}_X$ are of the form $h_j h_{n-j}$.

d. This result is equivalent to Exercise 7.47(j) in the special case that P is also $(\mathbf{2}+\mathbf{2})$-free. (It is also a special case of Exercise 7.92(a).) See R. Stanley and J. R. Stembridge, *J. Combinatorial Theory (A)* **62** (1993), 261–279 (§5). The weaker result that $\tilde{Z}_X$ is s-positive follows from Exercise 7.47(h) and also from Exercise 7.92(a).

e. In equation (7.186) put $k = 3$, $x^{(1)} = x$, $x^{(2)} = y$, $p_1(x^{(3)}) = 0$, $p_2(x^{(3)}) = p_3(x^{(3)}) = \cdots = 1$, $w_1 = v$, and $w_2 = w$. Using equation (7.209) we get

$$\sum_{\lambda \vdash n} H_\lambda \frac{d_\lambda}{n!} s_\lambda(x) s_\lambda(y) = \frac{1}{n!} \sum_{vw \in \mathfrak{D}_n} p_{\rho(v)}(x) p_{\rho(w)}(y).$$

Hence (since $\langle p_\mu, p_\nu \rangle = z_\mu \delta_{\mu\nu}$),

$$\tilde{Z}_{B_w}(x) = \left\langle \frac{1}{n!} \sum_{\lambda \vdash n} H_\lambda d_\lambda s_\lambda(x) s_\lambda(y),\ p_{\rho(w)}(y) \right\rangle_y ,$$

where $\langle\ ,\ \rangle_y$ indicates that we are taking the scalar product with respect to the y variables only. Since $p_\alpha = \sum_\lambda \chi^\lambda(\alpha) s_\lambda$ (Corollary 7.17.4) and $f^\lambda = n!/H_\lambda$ (Corollary 7.21.6), we get

$$\tilde{Z}_{B_w} = \sum_{\lambda \vdash n} (f^\lambda)^{-1} d_\lambda \chi^\lambda(w) s_\lambda,$$

as desired. This formula is a result of S. Okazaki, Ph.D. thesis, Massachusetts Institute of Technology, 1992 (Thm. 1.2).

f. Follows easily from (e) and Exercise 7.63(b). See Okazaki, *ibid.* (Thm. 1.6). Note that this result implies that $\tilde{Z}_{B_w}$ is s-positive when w is an n-cycle. In general, $\tilde{Z}_{B_w}$ need not be s-positive, e.g., if $w = \mathrm{id} \in \mathfrak{S}_2$, then $\tilde{Z}_{B_w} = s_2 - s_{11}$.

7.112. a. Let G be the subgroup of $\mathfrak{S}_n$ generated by the n-cycle $(1, 2, \ldots, n)$. A necklace with beads from an alphabet A is just an orbit of the action of G on $A^{[n]}$, the set of functions $[n] \to A$. Hence by Corollary 7.24.6, we have

$$N(k, n) = \frac{1}{n} \sum_{w \in G} k^{c(w)}.$$

Since G has $\phi(d)$ elements of cycle type $\langle d^{n/d} \rangle$, the proof follows. The enumeration (7.197) of necklaces is due to P. A. MacMahon, *Proc. London Math. Soc.* **23** (1892), 305–313 (p. 308); in *Collected Papers* (G. E. Andrews, ed.), MIT Press, Cambridge, 1978, pp. 468–476, and is a precursor of Pólya's theory of enumeration under group action. MacMahon mentions that the enumeration of necklaces according to the number of beads of each color (and hence including (b) as a special case) had earlier been done by M. E. Jablonski and M. Moreau, independently.

b. By Theorem 7.24.4, we want the coefficient of $x_1^n x_2^n$ in

$$Z_G(p_1, p_2, \ldots) = \frac{1}{2n} \sum_{d \mid 2n} \phi(d) p_d^{2n/d}.$$

It is easy to see that this coefficient is just

$$\frac{1}{2n} \sum_{d \mid n} \phi(d) \binom{2n/d}{n/d}.$$

7.113. Let V be a p-element set, and let G be the group of permutations of $S = \binom{V}{2}$ induced by permutations of V, as in Example 7.24.2(b). By Example 7.24.3(b) we have

$$\sum_{i=0}^{\binom{p}{2}} g_i(p) q^i = Z_G(1, q).$$

By equation (7.120) we have

$$Z_G(1,q) = \sum_{\lambda \vdash \binom{p}{2}} a_\lambda s_\lambda(1,q),$$

where $a_\lambda \in \mathbb{N}$. By Exercise 7.75(c) each polynomial $s_\lambda(1,q)$ is symmetric and unimodal, with center of symmetry $\frac{1}{2}\binom{p}{2}$. Hence the same is true of $Z_G(1,q)$, and the proof follows.

This result (in a more general form that can be proved in the same way as above) is due to D. Livingstone and A. Wagner, *Math. Z.* **90** (1965), 393–403. For further references and ramifications, see R. Stanley, *Ann. New York Acad. Sci.* **576** (1989), 500–535 (esp. Thm. 10) and *Discrete Appl. Math.* **34** (1991), 241–277 (§3).

Index

Additional Errata and Addenda

* p. 74, Exercise 5.8(a). The stated formula for $T(n, k)$ fails for $n = 0$. Also, it makes more sense to define $T(0, 0) = 1$.

* p. 81, Exercise 5.24(d). A solution was found by the Cambridge Combinatorics and Coffee Club (February 2000).

* p. 124, Exercise 5.28. A bijective proof based on Prüfer codes is due to the Cambridge Combinatorics and Coffee Club (December 1999).

* p. 124, Exercise 5.29(b). Update the Pitman reference to *J. Combinatorial Theory* (*A*) **85** (1999), 165–193. Further results on P_n and related posets are given by D. N. Kozlov, *J. Combinatorial Theory* (*A*) **88** (1999), 112–122.

* p. 136, last line of Exercise 5.41(j). A solution different from the one above was given by S. C. Locke, *Amer. Math. Monthly* **106** (1999), 168.

* p. 143, Exercise 5.50(c). The paper of Postnikov and Stanley has appeared in *J. Combinatorial Theory* (*A*) **91** (2000), 544–597.

* p. 144, Exercise 5.5.3. The identity

$$4^n = \sum_{j=0}^{n} 2^{n-j} \binom{n+j}{j} \tag{1}$$

  follows immediately from "Banach's match box problem," an account of which appears for instance in W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1, second ed., Wiley, New York, 1957 (§5.8). This yields a simple bijective proof of (1).

* p. 151, Exercise 5.62 (b). David Callan observed (private communication) that there is a very simple combinatorial proof. Any matrix of the type being enumerated can be written *uniquely* in the form $P + 2Q$, where P and Q are permutation matrices. Conversely $P + 2Q$ is always of the type being enumerated, whence $f_3(n) = n!^2$.

* p. 212. For further details on the history of Catalan numbers, see P. J. Larcombe and P. D. C. Wilson, *Mathematics Today* **34** (1998), 114–117; P. J. Larcombe, *Mathematics Today* **35** (1999), 25, 89; and P. J. Larcombe, *Math. Spectrum* **32** (1999/2000), 5–7.

* p. 217, Exercise 6.2(a). It needs to be assumed that $F(0) = 0$; otherwise e.g. $F(x) = 1/2$ is a trivial counterexample.

* p. 231, Exercise 6.25 (i). This conjecture has been proved by M. Haiman, A geometric proof of the $n!$ and Macdonald positivity conjectures, preprint available electronically at http://math.ucsd.edu/~mhaiman.

* p. 232, Exercise 6.27(c). Robin Chapman has found an elegant argument that there always exists an integral orthonormal basis.

* p. 250, Exercise 6.4. A complete description of a field of generalized power series that forms an algebraic closure of $\mathbb{F}_p[[x]]$ is given by K. S. Kedlaya, The algebraic closure of the power series field in positive characteristic, *Proc. Amer. Math. Soc.*, to appear.

* pp. 261–262, Exercise 3.19(pp). A further reference on noncrossing partitions is the nice survey article R. Simion, *Discrete Math.* **217** (2000), 367–409.

* p. 264, Exercise 6.19(iii). It should be mentioned that the diagonals of the frieze patterns of Exercise 6.19(mmm) are precisely the sequences $1a_1a_2\cdots a_n1$ of the present exercise.

* p. 265, Exercise 6.19(lll), lines 3− to 2−. The paper of Postnikov and Stanley has appeared in *J. Combinatorial Theory* (*A*) **91** (2000), 544–597.

* p. 265, Exercise 6.19(mmm). A couple of additional references to frieze patterns are H. S. M. Coxeter, *Acta Arith.* **18** (1971), 297–310, and H. S. M. Coxeter and J. F. Rigby, in *The Lighter Side of Mathematics* (R. K. Guy and R. E. Woodrow, eds.), Mathematical Association of America, Washington, DC, 1994, pp. 15–27.

* p. 269, line 1−, to p. 270, line 1. The paper of Postnikov and Stanley has appeared in *J. Combinatorial Theory* (*A*) **91** (2000), 544–597.

* p. 272, end of Exercise 6.33(c). Yet another proof was given by J. H. Przytycki and A. S. Sikora, Polygon dissectons and Euler, Fuss, Kirkman and Cayley numbers, preprint available electronically at math.CO/9811086.

* p. 279, Exercise 6.56(c). In the paper N. Alon and E. Friedgut, *J. Combinatorial Theory* (*A*) **89** (2000), 133–140, it is shown that $A_v(n) < c^n\gamma^*(n)$, where $\gamma^*(n)$ is an extremely slow growing function related to the Ackermann hierarchy. The paper is available electronically at http://www.ma.huji.ac.il/~ehudf.

* p. 291, line 9−. In general it is not true that $\hat{\Lambda}_R = \hat{\Lambda} \otimes R$; one only has a natural surjection from the former onto the latter. Equality will hold if R is noetherian.

* p. 295, Figure 7-3. In the expansion of h_{41}, the coefficient of m_{41} should be 2.

* p. 399, line 7−. For additional information concerning Craige Schensted, see the webpage ea.ea.home.mindspring.com.

* p. 439, reference A1.13. An updated version of this paper of van Leeuwen, entitled "The Littlewood-Richardson rule, and related combinatorics," is available electronically at math.CO/9908099.

* p. 467, Exercise 7.55(b). Let $f(n)$ be the number of $\lambda \vdash n$ satisfying (7.177). Then $(f(1), f(2), \dots, f(30)) = (1, 1, 1, 2, 2, 7, 7, 10, 10, 34, 40, 53, 61, 103, 112, 143, 145, 369, 458, 579, 712, 938, 1127, 1383, 1638, 2308, 2754, 3334, 3925, 5092)$.

* p. 484, Exercise 7.101(b). As in (a), the plane partitions being counted have largest part at most m.

* p. 485, line 7. The five displayed tableaux should be rotated 180°.

* p. 504, line 10−. Update the Babson, et al., reference to *Topology* **38** (1999), 271–299.

* p. 514, Exercise 7.47(m), lines 1–3. Update the reference to R. Stanley, *Discrete Math.* **193** (1998), 267–286.

* p. 515, Exercise 7.48(g). Further generalizations of shuffle posets are considered by P. Hersh, Two generalizations of posets of shuffles, preprint available electronically at http://www.math.washington.edu/~hersh/papers.html.

* p. 534, end of Exercise 7.74. For some connections between inner plethysm and graphical enumeration, see L. Travis, Ph.D. thesis, Brandeis University, 1999; available electronically at math.CO/9811127.

* p. 539, Exercise 7.85. A further reference to the evaluation of $g_{\lambda\mu\nu}$ is M. H. Rosas, The Kronecker product of Schur functions indexed by two-row shapes or hook shapes, preprint available electronically at math.CO/0001084.

* p. 544, lines 4− to 2−. Update the reference to R. Stanley, *Discrete Math.* **193** (1998), 267–286.

* p. 542, line 10. Update the Babson, et al., reference to *Topology* **38** (1999), 271–299.

* p. 551, Exercise 7.102(b), lines 2− to 1−. The "nice" bijective proof asked for was given by M. Rubey, A nice bijection for a content formula for skew semistandard Young tableaux, preprint. The proof is based on jeu de taquin.

* p. 554, last two lines of Exercise 7.107(a). Update reference to *Annals of Combinatorics* **2** (1998), 103–110.